Astérisque

JEAN-YVES CHEMIN

Explosion géométrique pour certaines équations d'ondes non linéaires

Astérisque, tome 266 (2000), Séminaire Bourbaki, exp. n° 850, p. 7-20
<http://www.numdam.org/item?id=SB_1998-1999__41__7_0>

© Société mathématique de France, 2000, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (http://smf4.emath.fr/Publications/Asterisque/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM
Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Séminaire BOURBAKI
51ème année, 1998-99, n° 850

Novembre 1998

EXPLOSION GÉOMÉTRIQUE POUR CERTAINES ÉQUATIONS D'ONDES NON LINÉAIRES
[d'après Serge Alinhac]

par **Jean-Yves CHEMIN**

1. INTRODUCTION

Le problème de l'apparition de singularités dans les équations aux dérivées partielles non linéaires est le problème qui domine leur étude. Que ce soit dans l'étude des systèmes d'Euler ou de Navier-Stokes (compressible ou incompressible), de l'équation de Schrödinger non linéaire ou des équations d'ondes non linéaires, on ne peut manquer d'être confronté à ce problème. La littérature consacrée à ce sujet est immense.

Dans ce texte, nous allons exposer les progrès récemment effectués dans la compréhension de l'apparition de singularités dans les équations d'ondes non linéaires. Commençons par préciser le cadre dans lequel nous allons travailler. Nous nous restreindrons au cas modèle de l'équation suivante, dans l'espace $\mathbf{R}^{1+d}$,

$$(E)\begin{cases} \partial_t^2 u - \Delta u - \displaystyle\sum_{0\le j,k\le d} g^{j,k}(\nabla u)\partial_j\partial_k u &= 0 \\ (u,\partial_t u)_{|t=0} &= \epsilon(u_0,u_1) \end{cases}$$

où les fonctions u_0 et u_1 sont des fonctions indéfiniment différentiables à support dans la boule de centre 0 et de rayon M, et ϵ est un (petit) paramètre strictement positif. On supposera de plus que les fonctions $g^{j,k}$ sont des fonctions linéaires de leurs arguments ; la notation ∇u désignant bien sûr le vecteur $(\partial_t u, \partial_{x_1} u \cdots, \partial_{x_d} u)$. Le terme

$$\sum_{1\le j,k\le d} g^{j,k}(\nabla u)\partial_j\partial_k u$$

doit être compris comme le premier terme d'un développement limité, la pertinence de ce point de vue est assurée par le fait que l'on regarde de petites données initiales.

2. MÉTHODE D'ÉNERGIE ET MINORATION DU TEMPS DE VIE

Il est classique que l'équation (E) est localement bien posée pour des données (u_0, u_1) appartenant à l'espace $H^s \times H^{s-1}$ pour s strictement supérieur à $d/2+2$. En suivant par exemple [23], on démontre, en utilisant des estimations d'énergie que, si u est solution de l'équation (E), alors

$$\frac{d}{dt}|\nabla u(t)|^2_{H^{s-1}} \leq C\|\nabla^2 u(t)\|_{L^\infty}|\nabla u(t)|^2_{H^{s-1}}. \tag{1}$$

L'inclusion de Sobolev assure que l'on a

$$\frac{d}{dt}|\nabla u(t)|^2_{H^{s-1}} \leq C|\nabla u(t)|^3_{H^{s-1}}.$$

Par intégration, il vient que

$$T \leq \frac{C}{|\nabla u(0)|_{H^{s-1}}} \Longrightarrow \sup_{t\in[0,T]} |\nabla u(t)|^2_{H^{s-1}} \leq 2|\nabla u(0)|^2_{H^{s-1}}.$$

Ainsi donc, le temps maximal d'existence T_ϵ est minoré par $C\epsilon^{-1}$. De plus, l'inégalité (1) implique que

$$|\nabla u(t)|^2_{H^{s-1}} \leq C|\nabla u(0)|^2_{H^{s-1}} \exp\Big(\int_0^t \|\nabla^2 u(\tau)\|_{L^\infty} d\tau\Big).$$

On en déduit alors que

$$T_\epsilon < \infty \Longrightarrow \int_0^{T_\epsilon} \|\nabla^2 u(\tau)\|_{L^\infty} d\tau = \infty.$$

Remarque.– L'équation (E) possède une propriété d'invariance par changement d'échelle, c'est-à-dire que, si u est solution de (E) sur un intervalle de temps $[0,T]$, alors la fonction

$$u_\lambda(t,x) \stackrel{\text{déf}}{=} \frac{1}{\lambda} u(\lambda t, \lambda x)$$

est solution de (E) sur l'intervalle $[0, T\lambda^{-1}]$.

Les méthodes d'énergie exposées ci-dessus font fi des propriétés particulières de l'équation des ondes. C'est sous l'impulsion de S. Klainermann que ces propriétés ont été utilisées de manière systématique pour améliorer le temps d'existence de solutions régulières notamment dans les travaux de S. Klainerman (voir [20]–[22]), et aussi F. John (voir [17]–[18]), F. John et S. Klainerman (voir [19]) et L. Hömander (voir [15]).

La propriété utilisée par S. Klainermann est l'invariance de l'équation des ondes sous l'action du groupe de Lorentz. Ceci l'a conduit à introduire les champs de vecteurs suivants.

DÉFINITION 2.1. — On désigne par $\mathcal{Z}$ la famille de champs de vecteurs suivants :

$$Z_j \stackrel{\text{déf}}{=} \partial_j\,,\ Z_{0,j} \stackrel{\text{déf}}{=} x_j\partial_t + t\partial_j \quad \text{et, pour } 1 \leq j < k \leq d, \quad Z_{j,k} \stackrel{\text{déf}}{=} x_j\partial_k - x_k\partial_j$$

ainsi que du champ radial $\rho \overset{\text{déf}}{=} t\partial_t + (x|\partial)$. De plus, si k est un entier positif, $\mathcal{Z}^k u$ désigne la famille formée par u et ses dérivées d'ordre au plus k par rapport aux éléments de la famille $\mathcal{Z}$.

Remarques.– Les champs de vecteurs $(Z_{0,j}, Z_{j,k}, \rho)$ forment une famille génératrice du C^∞-module des champs de vecteurs indéfiniment différentiables et tangents au cône

$$\Gamma \overset{\text{déf}}{=} \{(t,x) \in \mathbf{R}^{1+d} \,/\, t^2 - |x|^2 = 0\}.$$

De plus, ces champs de vecteurs vérifient les propriétés de commutation suivantes

$$[Z_{0,j}, \Box] = [Z_{j,k}, \Box] = 0 \quad \text{et} \quad [\rho, \Box] = 2\Box. \tag{2}$$

L'intérêt de ces champs de vecteurs réside pour nous dans l'estimation de Sobolev à poids suivante.

PROPOSITION 2.1. — *Il existe une constante C telle que, pour toute fonction régulière u, on ait, pour tout point (t,x) de $\mathbf{R}^{1+d}$,*

$$|(1+|t|+|x|)^{d-1}(1+||t|-|x||)\nabla u(t,x)|^2 \leq C\|\mathcal{Z}^{[\frac{d}{2}]+1}u(t,\cdot)\|^2_{\dot{H}^1},$$

où $\|a(t,\cdot)\|^2_{\dot{H}^1} \overset{\text{déf}}{=} \|\nabla a(t,\cdot)\|^2_{L^2}$.

Pour simplifier la démonstration de cette proposition, nous supposerons que $d = 2$ et que le support de u est inclus dans $\mathcal{B} \overset{\text{déf}}{=} \{(t,x) \in \mathbf{R}^3 \,/\, |x| \leq M+t\}$. Pour démontrer cette estimation, commençons par observer que, si la fonction u est à support compact en (t,x), alors il s'agit de l'estimation de Sobolev usuelle. On peut donc supposer que $|t|+|x| \geq 1$. Il convient alors de distinguer deux zones.

Supposons tout d'abord que le support de la fonction u est inclus dans l'ensemble des (t,x) tels que $|x| \notin \left[\frac{1}{2}t, \frac{3}{2}t\right]$. Soit a une fonction indéfiniment différentiable à support dans cet ensemble. L'inégalité de Sobolev classique implique que

$$|(t^2-|x|^2)a(t,x)| \leq Ct\|\partial_1\partial_2((t^2-|x|^2)a(t,x))\|_{L^2}.$$

Le fait que

$$\partial_j = \frac{t}{(t^2-|x|^2)}Z_{0,j} - \frac{x_j}{(t^2-|x|^2)}\rho + \sum_{k=1}^{d}\frac{x_k}{(t^2-|x|^2)}Z_{j,k}$$

implique que

$$\partial_1\partial_2((t^2-|x|^2)a(t,x)) = \sum_{|I|\leq 2}\alpha_I Z^I a\,;$$

où les fonctions α_I sont indéfiniment différentiables homogènes de degré 0. Il en résulte que

$$|(t^2-|x|^2)a(t,x)| \leq Ct\|\mathcal{Z}^2 a(t)\|_{L^2},$$

et donc comme $|x| \notin \left[\frac{1}{2}t, \frac{3}{2}t\right]$, on a

$$|(t^2 - |x|^2)^{\frac{1}{2}} a(t,x)| \leq C\|\mathcal{Z}^2 a(t)\|_{L^2}. \tag{3}$$

Considérons maintenant une fonction a à support dans $\left\{(t,x) \,/\, |x| \in \left[\frac{1}{2}t, \frac{3}{2}t\right]\right\}$. On redresse alors le cône Γ en posant

$$b(t,\sigma,\omega) \stackrel{\text{déf}}{=} a(t,x) \quad \text{où} \quad \sigma \stackrel{\text{déf}}{=} r - t$$

et où (r,ω) désigne les coordonnées polaires. Un calcul des plus élémentaires montre que

$$\begin{aligned} \rho a &= (t\partial_t + \sigma\partial_\sigma)b \quad \text{et} \\ \sum_{k=1}^{2} \frac{x_k}{|x|} Z_{0,k} a &= ((t+\sigma)\partial_t - \sigma\partial_\sigma)b. \end{aligned}$$

Il en résulte que

$$\sigma\partial_\sigma = \sum_{Z \in \mathcal{Z}} \alpha_Z Z \tag{4}$$

où les α_Z désignent des fonctions de (t,σ,ω) bornées ainsi que toutes leurs dérivées. Soit φ une partition de l'unité dyadique sur $\mathbf{R}$. L'inclusion de Sobolev classique implique que

$$|\varphi(2^{-p}\sigma) b(t,\sigma,\omega)| \leq C 2^{\frac{p}{2}} \|\partial_q \partial_\omega (\varphi(2^{-p}\cdot) b)\|_{L^2}.$$

Vu la définition de φ, on peut alors écrire

$$\begin{aligned} t^{\frac{1}{2}} 2^{\frac{p}{2}} |\varphi(2^{-p}\sigma) b(t,\sigma,\omega)| &\leq C t^{\frac{1}{2}} 2^p \|\partial_\sigma \partial_\omega (\varphi(2^{-p}\cdot) b)\|_{L^2} \\ &\leq C t^{\frac{1}{2}} \|\sigma \partial_\sigma \partial_\omega (\varphi(2^{-p}\cdot) b)\|_{L^2} \\ &\leq C t^{\frac{1}{2}} (\|\sigma \partial_\sigma \varphi(2^{-p}\cdot) \partial_\omega b\|_{L^2} + \|\varphi(2^{-p}\cdot) \sigma \partial_\sigma \partial_\omega b\|_{L^2}) \\ &\leq C t^{\frac{1}{2}} (\|\partial_\omega b\|_{L^2} + \|\sigma \partial_\sigma \partial_\omega b\|_{L^2}). \end{aligned}$$

La relation (4) implique que

$$t^{\frac{1}{2}} |\varphi(2^{-p}\sigma)| \sigma|^{\frac{1}{2}} b(t,\sigma,\omega)| \leq C\|\mathcal{Z}^2 a(t,\cdot)\|_{L^2}.$$

Lorsque $|\sigma| \leq 1$, il suffit d'observer que

$$\partial_\sigma = \frac{1}{r}\rho - \frac{t}{r}\partial_t\,,$$

puis d'appliquer l'inégalité de Sobolev standard. On applique alors l'inégalité ci-dessus et l'inégalité (3) à la fonction ∇u et l'on trouve que

$$|(1 + |t| + |x|)^{d-1} (1 + ||t| - |x||) \nabla u(t,x)|^2 \leq C\|\mathcal{Z}^{[\frac{d}{2}]+1} \nabla u(t,\cdot)\|_{L^2}^2.$$

Comme, pour tout élément Z de la famille $\mathcal{Z}$, le commutateur $[Z, \partial_j]$ vaut ∂_k ou bien 0, la proposition est démontrée.

Nous sommes maintenant en mesure de démontrer le théorème suivant.

THÉORÈME 2.2. — *On a les minorations suivantes pour le temps d'existence maximal* T_ϵ. *Si* $d \geq 4$, *alors* $T_\epsilon = +\infty$, *si* $d = 3$, $T_\epsilon \geq \exp\left(\frac{C}{\epsilon}\right)$ *et si* $d = 2$, $T_\epsilon \geq \frac{C}{\epsilon^2}$.

Pour démontrer ce théorème, nous procédons par estimation d'énergie dans les espaces de Sobolev construits à l'aide des champs de vecteurs de la famille $\mathcal{Z}$. Soit s un entier supérieur ou égal à $d+5$. On pose alors

$$\|\mathcal{Z}^s u(t,\cdot)\|^2_{\dot{H}^1} \overset{\text{déf}}{=} \sum_{|I|\leq s} \|\nabla \mathcal{Z}^I u(t,\cdot)\|^2_{L^2}.$$

La formule de Leibnitz et les propriétés de commutation (2) impliquent que

$$\Box \mathcal{Z}^I u = \sum_{0\leq j,k\leq d} g^{j,k}(\nabla u)\partial_j\partial_k \mathcal{Z}^I u = F_I \tag{5}$$

où F_I est de la forme

$$F_I = \sum_{\substack{J+K=I \\ J\neq 0}} \nabla(\mathcal{Z}^J u)\nabla^2(\mathcal{Z}^K u).$$

Estimons la norme L^2 de F_I. Comme $J+K=I$, ou bien $|J| \leq s/2$, ou bien $|K| < s/2$. Dans le premier cas, on écrit que

$$\|\nabla(\mathcal{Z}^J u(t,\cdot))\nabla^2(\mathcal{Z}^K u(t,\cdot))\|_{L^2} \leq \|\nabla(\mathcal{Z}^J u(t,\cdot))\|_{L^\infty}\|\nabla^2(\mathcal{Z}^K u(t,\cdot))\|_{L^2}.$$

Comme $J \neq 0$ et que $s \geq d+5$, il vient, en appliquant la proposition 2.1 que

$$\|\nabla(\mathcal{Z}^J u(t,\cdot))\nabla^2(\mathcal{Z}^K u(t,\cdot))\|_{L^2} \leq \frac{C}{(1+t)^{\frac{d-1}{2}}}\|\mathcal{Z}^s u(t,\cdot)\|^2_{\dot{H}^1}.$$

Si maintenant $|K| < s/2$, l'estimation est analogue ; d'où l'inégalité

$$\|F_I(t,\cdot)\|_{L^2} \leq \frac{C}{(1+t)^{\frac{d-1}{2}}}\|\mathcal{Z}^s u(t,\cdot)\|^2_{\dot{H}^1}. \tag{6}$$

L'estimation d'énergie usuelle assure que

$$\frac{d}{dt}\left(\|\mathcal{Z}^s u(t,\cdot)\|^2_{\dot{H}^1} \exp - \int_0^t \|\nabla^2 u(\tau,\cdot)\|_{L^\infty} d\tau\right) \leq \frac{C}{(1+t)^{\frac{d-1}{2}}}\|\mathcal{Z}^s u(t,\cdot)\|^3_{\dot{H}^1}.$$

Considérons de manière classique le réel $\underline{T}_\epsilon$ défini par

$$\underline{T}_\epsilon \overset{\text{déf}}{=} \sup\Big\{T < T_\epsilon \,/\, \sup_{t\leq T}\|\mathcal{Z}^s u(t,\cdot)\|_{\dot{H}^1} \leq 4\epsilon\|\mathcal{Z}^s u(0,\cdot)\|_{\dot{H}^1}\Big\}.$$

La proposition 2.1 implique alors que, pour tout $T < \underline{T}_\epsilon$, on a, si c est assez petit,

$$M^2_s(T) \overset{\text{déf}}{=} \sup_{t\leq T}\|\mathcal{Z}^s u(t,\cdot)\|^2_{\dot{H}^1} \leq \left(\epsilon\|\mathcal{Z}^s u(0,\cdot)\|^2_{\dot{H}^1} + CM^3_s(T)\int_0^t \frac{d\tau}{(1+\tau)^{\frac{d-1}{2}}}\right) \times \exp\left(C\epsilon\|\mathcal{Z}^s u(0,\cdot)\|_{\dot{H}^1}\int_0^t \frac{d\tau}{(1+\tau)^{\frac{d-1}{2}}}\right).$$

Le théorème en résulte immédiatement.

Remarques.– Cette méthode n'est pas économe en régularité sur les données initiales. Elle est basée en grande partie sur un effet dispersif de l'équation des ondes. Une autre façon d'utiliser cela consiste à s'appuyer sur des effets liés à l'estimation de Strichartz. On peut démontrer, en suivant par exemple la démarche de Ginibre et Velo (voir par exemple [14]) que les solutions de $\Box u = 0$ vérifient

$$\|\nabla u(t,\cdot)\|_{L^\infty} \leq \frac{C}{t^{\frac{d-1}{2}}}\|\nabla u_{|t=0}\|_B$$

où B est un espace de Besov du type "$d-2$ dérivées dans L^1". Malheureusement, cette inégalité ne peut être utilisée, car les normes de type L^1 ne sauraient être propagées par des équations de type onde.

Néammoins, en utilisant des effets de type Strichartz, on peut démontrer (voir [13] et [24]) qu'en dimension deux d'espace, si $(u_0,u_1) \in H^s \times H^{s-1}$ avec $s > \frac{23}{8}$, alors

$$T_\epsilon \geq \frac{C(\|(u_0,u_1)\|_{H^s\times H^{s-1}}}{\epsilon^{\frac{8}{7}}}.$$

3. UN CAS MODÈLE TRÈS SIMPLE D'APPARITION DE SINGULARITÉS : L'ÉQUATION DE BURGER'S

Il s'agit maintenant d'être capable de prédire si des singularités peuvent effectivement apparaître, c'est-à-dire si le théorème 2.2 ci-dessus est optimal. Il est un cas modèle très simple où l'on sait prédire exactement quand la singularité apparaît et quelle est la nature de celle-ci ; c'est l'équation modèle dite de Burger's

$$\partial_t u + u\partial_x u = 0 \quad \text{et} \quad u_{|t=0} = u_0$$

dans $\mathbf{R}_+ \times \mathbf{R}$.

Ce modèle a le mérite d'être tout à la fois très simple, de fournir un modèle d'explosion qui sera transposable au cas de l'équation des ondes quasilinéaire (E), et d'intervenir techniquement dans la démonstration du théorème 4.1.

La très classique méthode des caractéristiques nous dit que le graphe de la solution est l'ensemble

$$\mathcal{G} \stackrel{\text{déf}}{=} \{(t, x + tu_0(x), u_0(x))\,,\ (t,x) \in \mathbf{R}^+ \times \mathbf{R}\},$$

tant que cet ensemble est un graphe, c'est-à-dire tant que $1 + tu_0'(x)$ reste strictement positif. Si par exemple la donnée initiale u_0 est de classe C^2 à support compact, il existe nécessairement un tel point. Le temps maximal d'existence d'une solution régulière est alors donné par

$$T = \min -\frac{1}{u_0'(x)}.$$

De plus, en dérivant l'équation, on trouve que

$$\partial_t \partial_x u + u \partial_x^2 u + (\partial_x u)^2 = 0,$$

ce qui assure que, par intégration le long des lignes de flot :

$$\partial_x u(t, x + t u_0(x)) = \frac{\partial_x u_0(x)}{1 + t \partial_x u_0(x)}.$$

Soit x_0 un point tel que

$$T = -\frac{1}{u_0'(x_0)}.$$

On voit immédiatement que

$$\partial_x u(t, x_0 + t u_0(x_0)) = \frac{T}{T - t}. \tag{7}$$

La solution u est constante sur les courbes caractéristiques, ce qui peut s'écrire

$$u(t, \phi(t, x)) = v(t, x)$$

où ϕ et v vérifient les sytèmes

$$\left\{ \begin{array}{rcl} \partial_t \phi(t,x) & = & u_0(x) \\ \phi(0,x) & = & x \end{array} \right. \quad \text{et} \quad \left\{ \begin{array}{rcl} \partial_t v(t,x) & = & 0 \\ v(0,x) & = & u_0(x). \end{array} \right. \tag{8}$$

Les singularités de la solution correspondent aux points où $\partial_x \phi$ s'annule sans que $\partial_x u_0$ ne s'annule.

Supposons que x_0 soit l'unique minimum non dégénéré de u_0' (la valeur de u_0' en ce point est alors strictement négative puisque u est supposée à support compact). Alors la fonction ϕ vérifie, pour tout $t \leq T$,

$$(H_0) \left\{ \begin{array}{l} \partial_x \phi(t,x) \geq 0 \quad \text{et} \quad \partial_x \phi(t,x) = 0 \Longleftrightarrow (t,x) = (T, x_0), \\ \\ \partial_x \partial_t \phi(T, x_0) < 0\,, \ \partial_x^2 \phi(T, x_0) = 0 \quad \text{et} \quad \partial_x^3 \phi(T, x_0) > 0. \end{array} \right.$$

Une telle singularité est une singularité de type cusp.

Cette vision du problème est à la base du concept d'explosion géométrique (et de la démonstration du théorème d'explosion 4.1).

4. ÉNONCÉ DU THÉORÈME PRINCIPAL

Dans toute la suite, nous supposerons que la dimension d vaut 2. Le théorème 4.1 ci-dessous montre que l'ordre de grandeur du temps d'existence T_ϵ obtenu par les méthodes d'énergie ci-dessus est le bon.

Pour simplifier, on supposera que l'équation (E) est de la forme

$$(E) \left\{ \begin{array}{rcl} \partial_t^2 u - \Delta u - \partial_t u \partial_t^2 u & = & 0 \\ (u, \partial_t u)_{|t=0} & = & \epsilon(u_0, u_1) \end{array} \right.$$

Avant d'énoncer le théorème, définissons la quantité qui gouverne le temps d'existence.

DÉFINITION 4.1. — On appelle premier profil associé aux données intiales (u_0, u_1), et l'on désigne par $R^{(1)}$ la fonction définie par

$$R^{(1)}(\sigma,\omega) \overset{\text{déf}}{=} \frac{1}{2\sqrt{2\pi}} \int_{s\geq\sigma} \frac{1}{\sqrt{s-\sigma}} (R(s,\omega,u_1) - \partial_s R(s,\omega,u_0))ds,$$

où R désigne la transformation de Radon

$$R(s,\omega,v) \overset{\text{déf}}{=} \int_{(x|\omega)=s} v(x)dx.$$

Le théorème qui suit précise non seulement le temps d'existence, mais estime la solution près du point d'explosion.

THÉORÈME 4.1. — *Supposons que la fonction $\partial_\sigma^2 R^{(1)}$ ait un unique minimum négatif et non dégénéré en (σ_0,ω_0) et posons*

$$\tau_0 \overset{déf}{=} -\frac{1}{\partial_\sigma^2 R^{(1)}(\sigma_0,\omega_0)}.$$

- *Le temps d'existence T_ϵ satisfait*

$$T_\epsilon = \frac{\tau_0}{\epsilon^2} + O\left(\frac{1}{\epsilon}\right).$$

- *De plus, il existe une constante C (dépendant de (u_0,u_1)) et une famille $(x_\epsilon)_{0<\epsilon<\epsilon_0}$ telle que $||x_\epsilon| - T_\epsilon| \leq C$ et telle que, pour $\tau_0' < \tau_0$, on ait, en posant $I_\epsilon \overset{déf}{=} [\tau_0'\epsilon^{-2}, T_\epsilon]$,*

$$\|\nabla u\|_{L^\infty(I_\epsilon\times\mathbf{R}^2)} \leq C\epsilon^2. \tag{9}$$

- *Pour tout α strictement positif, il existe C_α tel que*

$$\|\nabla^2 u\|_{L^\infty((I_\epsilon\times\mathbf{R}^2)\cap\mathcal{D}_\epsilon^\alpha)} \leq C\epsilon^2 \quad avec \quad \mathcal{D}_\epsilon^\alpha \overset{déf}{=} \mathbf{R}^{1+2}\setminus B((T_\epsilon,x_\epsilon),\alpha\epsilon^{-2}). \tag{10}$$

- *Enfin, on a*

$$\|\nabla^2 u(t,\cdot)\|_{L^\infty} \leq \frac{C}{T_\epsilon - t} \quad et \quad \|\partial_t^2 u(t,\cdot)\|_{L^\infty} \geq \frac{1}{C(T_\epsilon - t)}. \tag{11}$$

Remarques.– Les inégalités (11) nient de manière "critique" l'appartenance de $\nabla^2 u$ à l'espace $L^1([0,T_\epsilon];L^\infty)$. Elles ont une analogie frappante avec l'inégalité (7).

De plus, la quantité $\sup_{t\in[0,T_\epsilon[} (T_\epsilon - t)\|\nabla u(t,\cdot)\|_{L^\infty}$ est invariante par le changement d'échelle de l'équation.

La démonstration de ce résultat est l'aboutissement d'une longue série d'articles de S. Alinhac (les références [1] à [12]). La très grande technicité de la preuve nous interdit bien sûr d'en donner le détail. Nous nous contenterons d'en expliquer les grandes lignes et les idées fortes.

Tout d'abord, il convient de faire une analyse asymptotique par rapport au paramètre ϵ pour dégager l'allure de la solution bien avant l'apparition de la singularité. Cette analyse

asymptotique permet alors de ramener le problème à la résolution d'une équation aux dérivées partielles du second ordre (dépendant de ϵ), la solution construite présentant une singularité en un point. La résolution de cette équation aux dérivées partielles se fait par une procédure d'éclatement.

5. ANALYSE ASYMPTOTIQUE ET OPTIQUE GÉOMÉTRIQUE

L'analyse asymptotique qui sert de base à la démonstration est faite dans [3] (on pourra aussi consulter [15]). Nous allons exposer ici de manière heuristique les idées importantes de la preuve.

Tout d'abord, les estimations de Sobolev à poids de la proposition 2.1 nous disent qu'à l'intérieur du cône de lumière, c'est-à-dire dans la zone où $t - r$ est assez grand, la quantité $\nabla^2 u$ est mieux contrôlée, et donc que les singularités ne sauraient apparaître dans cette zone. Il convient donc de se concentrer sur le bord du cône de lumière, c'est-à-dire dans la zone où $-C \leq r - t \leq M$.

Dans cette zone, il convient de s'intéresser particulièrement à l'allure de la solution du problème linéaire

$$\begin{cases} \Box u^{(1)} &= 0 \\ (u, \partial_t u) &= \epsilon(u_0, u_1). \end{cases}$$

Il est bien connu (voir par exemple [16]) qu'elle peut s'écrire sous la forme

$$u^{(1)} = \frac{\epsilon}{r^{\frac{1}{2}}} F\big(r - t, \omega, \frac{1}{r}\big).$$

Les propriétés de l'équation des ondes en dimension 2 d'espace n'impliquent pas que F soit à support compact, mais simplement que F vérifie des propriétés de type symbole, c'est-à-dire que

$$|\partial_\sigma^\alpha \partial_\omega^\beta \partial_z^\gamma F(\sigma, \omega, z)| \leq C_{\alpha,\beta,\gamma}(1 + |\sigma|)^{-\frac{1}{2}+\gamma-\alpha}.$$

Une formule de Taylor en $z = 0$ implique que

$$u^{(1)} \sim \frac{\epsilon}{r^{\frac{1}{2}}}\left(R^{(1)}(r - t, \omega) + \frac{1}{r}R^{(2)}(r - t, \omega) \cdots + \frac{1}{r^k}R^{(k+1)}(r - t, \omega) + \cdots\right)$$

où les $R^{(k)}$ vérifient des propriétés de type symbole

$$|\partial_\sigma^\alpha \partial_\omega^\beta R^{(k+1)})(\sigma, \omega)| \leq C_{k,\alpha,\beta}(1 + |\sigma|)^{-\frac{1}{2}+k-\alpha}.$$

On en déduit donc que, au bord du cône, c'est-à-dire dans la zone où $-C \leq r - t \leq M$, la solution $u^{(1)}$ du problème linéaire est proche de

$$\frac{\epsilon}{r^{\frac{1}{2}}} R^{(1)}(r - t, \omega),$$

et ce, dès que t est assez grand, par exemple de l'ordre de ϵ^{-1}.

Suivant alors une idée d'optique géométrique, on cherche la solution u sous la forme d'un profil modulé, c'est-à-dire sous la forme

$$u_\epsilon = \frac{\epsilon}{r^{\frac{1}{2}}} G_\epsilon(r-t, \omega, \epsilon t^{\frac{1}{2}}),$$

la variable $\tau \overset{\text{déf}}{=} \epsilon t^{\frac{1}{2}}$ étant appelée "temps lent".

Un calcul immédiat montre que, si u est solution exacte de (E), alors la fonction G_ϵ est solution exacte de

$$\partial_\sigma \partial_\tau G_\epsilon + \partial_\sigma G_\epsilon (\partial_\sigma^2 G_\epsilon) + \epsilon \sum_{0 \leq j,k \leq d} P_\epsilon^{j,k}(\sigma, \omega, \tau, G_\epsilon, \nabla G_\epsilon) \partial_j \partial_k G_\epsilon = 0, \tag{12}$$

les $P^{j,k}$ étant des fonctions indéfiniment différentiables de leurs arguments (uniformément en ϵ) dans la zone où $-C \leq r - t \leq M$.

Les méthodes d'énergie utilisant la famille de champs de vecteurs $\mathcal{Z}$ permettent de penser que, tant que $t \leq \epsilon^{-2}\tau_0'$, avec $\tau_0' < \tau_0$ et près du bord du cône d'onde, la solution u se comporte comme

$$\frac{\epsilon}{r^{\frac{1}{2}}} G^{(1)}(r-t, \omega, \epsilon t^{\frac{1}{2}})$$

où $G^{(1)}$ est solution, non pas de l'équation (12), mais de l'équation

$$\begin{cases} \partial_\sigma \partial_\tau G^{(1)} + \partial_\sigma G^{(1)} (\partial_\sigma^2 G^{(1)}) &= 0 \\ \partial_\sigma G^{(1)}_{|\tau=0}(\sigma, \omega) &= R^{(1)}(\sigma, \omega). \end{cases}$$

Remarquons que le temps de vie de cette équation de Burger's en $\partial_\sigma G^{(1)}$ est exactement $-\dfrac{1}{\partial_\sigma^2 R^{(1)}(\sigma_0, \omega_0)}$. C'est ainsi qu'il est démontré dans [15] que

$$\liminf_{e \to 0} \epsilon^2 T_\epsilon \geq \tau_0 = -\frac{1}{\partial_\sigma^2 R^{(1)}(\sigma_0, \omega_0)}.$$

Ainsi, la bonne équation pour le profil modulé G_ϵ apparaît être

$$(EP_\epsilon) \begin{cases} \partial_\sigma \partial_\tau G_\epsilon + \partial_\sigma G_\epsilon (\partial_\sigma^2 G_\epsilon) &= -\epsilon^2 \sum_{0 \leq j,k \leq d} P_\epsilon^{j,k}(\sigma, \omega, \tau, G_\epsilon, \nabla G_\epsilon) \partial_j \partial_k G_\epsilon \\ \partial_\sigma G_{\epsilon|\tau=\tau_0'}(\sigma, \omega) &= G^{(1)}(\sigma, \omega, \tau_0') + \epsilon L_\epsilon(\sigma, \omega). \end{cases} \tag{13}$$

où L est une fonction indéfiniment différentiable (uniformément en ϵ).

6. LE THÉORÈME D'EXPLOSION GÉOMÉTRIQUE

Le théorème 4.1 est un corollaire du théorème plus précis suivant.

THÉORÈME 6.1. — *Pour tout entier k supérieur ou égal à 3, il existe un réel strictement positif ϵ_k et un réel strictement positif C_0 et un ouvert relativement compact V de $\mathbf{R} \times \mathbf{S}^1 \times \mathbf{R}$ tels que, pour tout $\epsilon \in]0, \epsilon_k[$, l'ensemble*

$$V_\epsilon \overset{\text{déf}}{=} V \cap (\mathbf{R} \times \mathbf{S}^1 \times]-\infty, \epsilon^2 T_\epsilon])$$

soit non vide. Il existe de plus trois familles (ϕ_ϵ), (v_ϵ) et (w_ϵ), uniformément de classe C^k sur V_ϵ et une famille $(\widetilde{M}_\epsilon)$ de points, avec $\widetilde{M}_\epsilon = (\sigma_\epsilon, \omega_\epsilon, \epsilon^2 T_\epsilon) \in V_\epsilon$, vérifiant les propriétés suivantes.

- *La fonction ϕ_ϵ vérifie*

$$(H)\begin{cases} \partial_s\phi_\epsilon \geq 0 \quad et \quad \partial_s\phi_\epsilon(s,\sigma,\omega) = 0 \Longleftrightarrow (s,\sigma,\omega) = \widetilde{M}_\epsilon \\ \\ \partial_s\partial_\tau\phi_\epsilon(\widetilde{M}_\epsilon) < 0\,,\ \partial_{s,\omega}\partial_s\phi_\epsilon(\widetilde{M}_\epsilon) = 0 \quad et \quad \partial^2_{s,\omega}\partial_s\phi_\epsilon(\widetilde{M}_\epsilon) \gg 0. \end{cases}$$

- *On a de plus*

$$\partial_s w_\epsilon = (\partial_s\phi)v_\epsilon \quad et \quad \partial_s v_\epsilon(\widetilde{M}_\epsilon) \neq 0.$$

- *Posons maintenant*

$$\begin{aligned} \Phi_\epsilon(s,\omega,\tau) &\overset{déf}{=} (\phi_\epsilon(s,\omega,\tau),\omega,\tau)\,, \\ x_\epsilon &\overset{déf}{=} (\sigma_\epsilon + T_\epsilon)\omega_\epsilon \quad et \\ \overline{M}_\epsilon &\overset{déf}{=} \left(|x_\epsilon| - T_\epsilon, \frac{x_\epsilon}{|x_\epsilon|}, \epsilon^2 T_\epsilon\right). \end{aligned}$$

Il existe alors un réel strictement positif α tel que

$$\forall (t,x) \in]-\infty, T_\epsilon] \cap B((T_\epsilon, x_\epsilon), \alpha\epsilon^{-2})\,,\ u = \frac{\epsilon}{r^{\frac{1}{2}}} G_\epsilon(r-t, \omega, \epsilon t^{\frac{1}{2}}),$$

la fonction G_ϵ étant définie sur un voisinage de $\overline{M}_\epsilon$ dans $\mathbf{R}\times\mathbf{S}^1\times]-\infty, \epsilon^2 T_\epsilon]$ par

$$G_\epsilon(\Phi_\epsilon) = w_\epsilon.$$

Remarque.– Les conditions (H) sont les analogues exactes des conditions (H_0) ; la singularité est de type cusp.

La démonstration de ce théorème n'est, compte tenu de l'analyse asymptotique précédemment exposée, rien d'autre que la résolution de l'équation (EP_ϵ).

7. L'ÉCLATEMENT D'UN SYSTÈME SUR UN EXEMPLE MODÈLE

Plutôt que de résoudre le système (EP_ϵ) général, nous allons exposer le cas modèle suivant :

$$\partial_\sigma\partial_\tau G + \partial_\sigma G(\partial^2_\sigma G) + \epsilon^2\partial^2_\omega G = 0.$$

Si l'on pose $G_1 \overset{\text{déf}}{=} \partial_\sigma G$ et $G_2 \overset{\text{déf}}{=} \partial_\omega G$, on trouve que l'équation ci-dessus est équivalente au système suivant

$$(EPM_\epsilon)\begin{cases} \partial_\tau G_1 + G_1\partial_\sigma G_1 + \epsilon^2\partial^2_\omega G_2 &= 0 \\ \partial_\sigma G_2 - \partial_\omega G_1 &= 0. \end{cases}$$

La procédure d'éclatement consiste à rechercher (G_1, G_2) sous la forme

$$(G_1, G_2)(\Phi) = (v_1, v_2) \quad \text{avec} \quad \Phi(X, Y, T) = (\sigma = \phi(X, Y, T), \omega = Y, \tau = T),$$

ce qui conduit au système

$$\left\{ \begin{array}{rcl} \partial_T \phi - v_1 - \epsilon^2 (\partial_Y \phi)^2 & = & 0 \\ \partial_T v_1 - \epsilon^2 \partial\phi \partial_Y v_1 + \epsilon^2 \partial_Y v_2 & = & 0 \\ \partial_X v_2 - \partial_X \phi \partial_Y v_1 + \partial_Y \phi \partial_X v_1 & = & 0 \end{array} \right.$$

avec les conditions aux limites

$$\begin{array}{rcl} \phi(X, Y, 0) & = & X \\ v_1(X, Y, 0) & = & \partial_X G \\ v_2(X, Y, 0) & = & \partial_Y G \quad \text{et} \end{array}$$

$$\phi(0, X, Y) = v_1(0, Y, Y) = v_2(0, Y, T) = 0.$$

Lorsque $\epsilon = 0$, le lecteur aura reconnu l'écriture "éclatée" de l'équation de Burger's donnée par (8). Lorsque ϵ est strictement positif, le système ci-dessus est l'éclaté du système (EPM_ϵ).

THÉORÈME 7.1. — *Sous l'hypothèse que $\partial_\sigma^2 G$ admet un unique minimum strictement négatif et non dégénéré, alors pour tout entier pair s assez grand et pour tout μ strictement supérieur à 1, il existe, pour ϵ assez petit, un temps $\widetilde{T}_\epsilon$ et un domaine V_ϵ de forme analogue à celui du théorème 6.1, tels qu'il existe trois fonctions ϕ_ϵ, $v_{1,\epsilon}$ et $v_{2,\epsilon}$ définies sur V_ϵ solutions du système ci-dessus. La fonction ϕ_ϵ vérifie :*

$$\sup_{\substack{\epsilon \\ j \leq 4}} \sum_{2k+\ell \leq s} \int_0^{\widetilde{T}_\epsilon} (\widetilde{T}_\epsilon - T)^\mu |\partial_X^k \partial_Y^\ell \partial_T^j \phi_\epsilon(X, Y, T)|^2 dX dY dT < \infty,$$

le système donnant les régularités correspondantes de $v_{1,\epsilon}$ et $v_{2,\epsilon}$.

De plus, ϕ_ϵ vérifie les conditions (H).

Les principales difficultés rencontrées dans la démonstration sont les suivantes. Tout d'abord, le linéarisé devient instable lorsque $\partial_X \phi$ devient négatif. Ceci empêche de résoudre le système au-delà du moment où $\partial_X \phi$ devient négatif. Mais ce moment est bien sûr une inconnue du problème. Il s'agit donc d'un problème à frontière libre.

La première étape consiste à se ramener à un problème sur un domaine fixe, ensuite on met en place une procédure à la Nash-Moser qui nécessite une étude détaillée de l'opérateur linéarisé.

Les détails sont donnés dans [9] pour le cas modèle étudié ici, dans [10] pour l'équation (EP_ϵ) et dans [11] pour l'équivalent de l'équation (EP_ϵ) dans le cas des équations (E) générales.

Références

[1] S. Alinhac - Une solution approchée en grands temps des équations d'Euler compressibles axisymétriques en dimension deux, *Communications in Partial Differential Equations* **17** (1992), 447–490.

[2] S. Alinhac - Temps de vie des solutions régulières des équations d'Euler compressibles axisymétriques en dimension deux, *Inventiones Mathematicae* (1993), 627–678.

[3] S. Alinhac - Approximation près du temps d'explosion des solutions d'équations d'onde quasi-linéaires en dimension deux, *SIAM Journal of Mathematical Analysis* **26** (1995), 529–565.

[4] S. Alinhac - Temps de vie précisé et explosion géométrique pour des systèmes hyberboliques quasilinéaires en dimension un d'espace I, *Annali della Scuola Normale Superiore di Pisa* **22** (1995), 493–515.

[5] S. Alinhac - Temps de vie précisé et explosion géométrique pour des systèmes hyberboliques quasilinéaires en dimension un d'espace II, *Duke Mathematical Journal* **73** (1994), 543–560.

[6] S. Alinhac - Explosion géométrique pour des systèmes quasilinéaires, *American Journal of Mathematics* **117** (1995), 987–1017.

[7] S. Alinhac - Temps de vie et comportement explosif des solutions d'équations d'ondes quasi-linéaires en dimension deux I, *Annales Scientifiques de l'École Normale Supérieure* **28** (1995), 225–251.

[8] S. Alinhac - Temps de vie et comportement explosif des solutions d'équations d'ondes quasi-linéaires en dimension deux II, *Duke Mathematical Journal* **73** (1994), 543–560.

[9] S. Alinhac - Explosion des solutions d'une équation d'ondes quasi-linéaire en deux dimensions d'espace, *Communications in Partial Differential Equations* **21** (1996), 923–969.

[10] S. Alinhac - Blow up of small data solutions for a class of quasilinear wave equations in two space dimensions I, *à paraître dans Annals of Mathematics*, 1998.

[11] S. Alinhac - Blow up of small data solutions for a class of quasilinear wave equations in two space dimensions II, *à paraître dans Acta Matematica.*

[12] S. Alinhac - Stability of geometric blowup, *Prépublication de l'Université Paris-Sud* 97–86.

[13] H. Bahouri et J.-Y. Chemin - Inégalités de Strichartz et équations d'ondes quasilinéaires, *Séminaire Équations aux dérivées Partielles de l'École Polytechnique*, 1997–1998.

[14] J. Ginibre et G. Velo - Generalized Strichartz inequalities for the wave equation, *Journal of Functional Analysis* **133** (1995), 50–68.

[15] L. Hörmander - The lifespan of classical solutions of non linear hyperbolic equations, *Lecture Notes in Mathematics* **1256**, Springer Verlag (1986), 214–280.

[16] L. Hörmander - *Lectures on Nonlinear Hyperbolic Differential Equations*, Mathématiques et Applications **26**, Springer, 1996.

[17] F. John - Blow-up of radial solutions of $u_{tt} = c^2(u_t)\Delta u$ in three space dimensions, Mat. Apl. Comput. **4** (1985), 3–18.

[18] F. John - Existence for large times of strict solutions of non linear wave equations in three space dimensions, *Communications in Pure and Applied Mathematics* **40** (1987), 79–109.

[19] F. John et S. Klainerman - Almost global existence to non linear wave equations in three space dimensions, *Communications in Pure and Applied Mathematics* **37** (1984), 443–455

[20] S. Klainerman - Global existence for nonlinear wave equations, *Communications in Pure and Applied Mathematics* **33** (1980), 43–101.

[21] S. Klainerman - Uniform decay estimates and the Lorentz invariance of the classical wave equation *Communications in Pure and Applied Mathematics* **38** (1985), 321–332.

[22] S. Klainerman - The null condition and global existence to non linear wave equations, *Communications in Pure and Applied Mathematics* **38** (1985), 631–641.

[23] A. Majda - *Compressible fluids flows and systems of conservation laws*, Springer Applied Mathematical Sciences **53** (1984).

[24] D. Tataru - Stichartz estimates for operators with nonsmooth coefficients and the nonlinear wave equation, *prépublication.*

[25] C. Zuily - Solutions en grand temps d'équations d'ondes non linéaires, in *Séminaire Bourbaki Volume 1993/1994*, Astérisque **227** (1995), 107–144.

Jean-Yves CHEMIN

Analyse Numérique, Case 187
Université Pierre et Marie CURIE
4 Place Jussieu
F-75230 Paris Cedex 05, France
Télécopie : (33) 01 44 27 72 00
E-mail : `chemin@ann.jussieu.fr`

Astérisque

PIERRE COLMEZ

Fonctions L p-adiques

Astérisque, tome 266 (2000), Séminaire Bourbaki, exp. n° 851, p. 21-58
<http://www.numdam.org/item?id=SB_1998-1999__41__21_0>

© Société mathématique de France, 2000, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (http://smf4.emath.fr/Publications/Asterisque/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM
Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Séminaire BOURBAKI Novembre 1998
51ème année, 1998-99, n° 851

FONCTIONS L p-ADIQUES

par **Pierre COLMEZ**

Introduction

On a à présent une bonne compréhension (conjecturale) du comportement aux entiers des fonctions L complexes des motifs (conjectures de Deligne-Beilinson et de Bloch-Kato [15, 6, 46, 21]; on pourra consulter le rapport Bourbaki de Fontaine [19] à ce sujet) et on aimerait bien obtenir le même genre de résultats pour leurs avatars p-adiques. Le premier problème auquel on est confronté est qu'on ne dispose pas (pour le moment) de définition p-adique de la fonction L p-adique d'un motif; la fonction L complexe est "définie" par un produit eulérien qui converge dans un demi-plan alors que ce même produit eulérien diverge p-adiquement. De toute façon, même sur les complexes la définition en termes de produit eulérien ne donne à peu près aucun résultat intéressant et pour démontrer quoi que ce soit sur une fonction L complexe, il faut commencer par en trouver une autre définition; de fait la seule méthode dont on dispose à l'heure actuelle est de commencer par relier la fonction L complexe à la théorie des formes automorphes.

Comme on ne sait pas définir p-adiquement les fonctions L p-adiques, on essaie de les construire par interpolation des valeurs aux entiers des fonctions L complexes. Par exemple, si $p \geq 3$, la fonction zêta p-adique (fonction zêta de Kubota-Leopoldt) a $p-1$ branches $\zeta_{p,i}$ pour $i \in \mathbf{Z}/(p-1)\mathbf{Z}$ et chacune de ces branches est construite en interpolant une partie des valeurs de la fonction zêta de Riemann aux entiers; de manière précise,

$$\zeta_{p,i}(n) = \begin{cases} (1-p^{-n})\zeta(n) & \text{si } i \text{ est impair et } n \leq 0 \text{ est congru à } i \text{ modulo } p-1, \\ \frac{\Gamma(n)}{(2i\pi)^n}(1-p^{n-1})\zeta(n) & \text{si } i \text{ est pair et } n \geq 1 \text{ est congru à } i \text{ modulo } p-1. \end{cases}$$

Comme $i+(p-1)\mathbf{N}$ et $i-(p-1)\mathbf{N}$ sont denses dans $\mathbf{Z}_p$, cela détermine $\zeta_{p,i}$ de manière unique, le miracle étant qu'il existe une fonction raisonnable prenant ces valeurs.

Dans l'exemple de la fonction zêta de Kubota-Leopoldt, les valeurs que l'on interpole sont, d'après Euler, des nombres algébriques (et même rationnels); on est dans le cas favorable où l'on dispose de beaucoup de valeurs critiques au sens de Deligne [15]. La plupart des fonctions L p-adiques que l'on sait construire présupposent un lien avec la

théorie des formes automorphes et l'existence de suffisamment de valeurs critiques [11] ; c'est le cas en particulier de celles construites par Manin [32] et Vishik [53], Amice et Vélu [3], Deligne et Ribet, Katz, Hida, Panchishkin... et nous renvoyons à [23, 24] pour ce point de vue que nous ne développerons pas, faute de place.

Dans cet exposé, nous allons décrire le point de vue de Perrin-Riou [41] qui est assez conjectural, mais fournit une image satisfaisante de ce que l'on peut espérer dans le cas de bonne réduction. En particulier, nous allons présenter une machine introduite par Perrin-Riou dans [40] qui permet de construire la fonction L p-adique d'un motif ayant bonne réduction en p à partir d'un "système compatible d'éléments spéciaux". Dans le cas de la fonction zêta de Kubota-Leopoldt, les éléments spéciaux sont les unités cyclotomiques et on retombe, en utilisant la machine de Perrin-Riou, sur la construction de la fonction zêta de Kubota-Leopoldt via les séries de Coleman [13], construction qui avait été inspirée par la construction par Coates et Wiles [12] de la fonction L p-adique d'une courbe elliptique à multiplication complexe à partir du système des unités elliptiques. Pour le moment, l'existence d'un "système compatible d'éléments spéciaux" relève, dans le cas général, de l'utopie mais, si on regarde les formules fournies par la machine de Perrin-Riou en ayant en tête le fait que les éléments spéciaux donnent, via une application "régulateur", les valeurs de la fonction L complexe, on obtient une conjecture sur les valeurs aux entiers de la fonction L p-adique que l'on peut voir soit comme l'analogue p-adique de la conjecture de Deligne-Beilinson, soit comme une définition (conjecturale) de la fonction L p-adique d'un motif par interpolation à partir des valeurs de sa fonction L complexe.

D'autre part, la machine de Perrin-Riou permet d'associer, sous des hypothèses beaucoup plus raisonnables que ci-dessus, une fonction L d'Iwasawa à un motif ayant bonne réduction en p. Cette fonction n'est bien définie qu'à multiplication près par une unité dans l'algèbre d'Iwasawa et on conjecture qu'elle ne diffère de la fonction L p-adique que par une unité dans l'algèbre d'Iwasawa (conjecture principale). Dans le cas du motif trivial, on retombe sur la conjecture d'Iwasawa reliant la fonction zêta de Kubota-Leopoldt aux groupes de classes d'idéaux des corps cyclotomiques, conjecture démontrée par Mazur et Wiles [34] et plus récemment par Rubin [44] en utilisant la méthode des systèmes d'Euler de Kolyvagin [31]. La conjecture principale peut aussi être énoncée en termes d'éléments spéciaux sans introduire de fonction L p-adique ; c'est le point de vue développé par Kato [28]. Comme la fonction L d'Iwasawa n'est bien définie qu'à une unité près dans l'algèbre d'Iwasawa, ses valeurs aux entiers ne sont bien définies qu'à multiplication près par une unité p-adique et le résultat principal de Perrin-Riou [41, chap. III] est que cette fonction L d'Iwasawa vérifie (l'analogue de) la conjecture de Bloch-Kato à une unité près, ce qui constitue le résultat général le plus encourageant en direction de la conjecture de Bloch-Kato. Signalons pour finir que la méthode des systèmes d'Euler [43, 29, 45] permet, dans le cas où la fonction L p-adique est obtenue à partir d'un système d'éléments spéciaux, de prouver une relation de divisibilité entre la fonction L d'Iwasawa et la fonction L p-adique.

1. EXEMPLES DE FONCTIONS L p-ADIQUES

1.1. Intégration sur $\mathbf{Z}_p$

1.1.1. *Quelques espaces fonctionnels.* Soient $\overline{\mathbf{Q}}_p$ une clôture algébrique de $\mathbf{Q}_p$ et $\mathbf{C}_p$ le complété de $\overline{\mathbf{Q}}_p$ pour la norme p-adique. Soit $\mathscr{C}^0$ l'espace des fonctions continues de $\mathbf{Z}_p$ dans $\mathbf{C}_p$. On fait de $\mathscr{C}^0$ un espace de Banach p-adique en le munissant de la norme $\| \ \|_{\mathscr{C}^0}$ définie par $\|f\|_{\mathscr{C}^0} = \sup_{x\in\mathbf{Z}_p} |f(x)|$.

Si $h \in \mathbf{N}$, soit LA_h l'espace des fonctions de $\mathbf{Z}_p$ dans $\mathbf{C}_p$ analytiques sur $a + p^h\mathbf{Z}_p$ quel que soit $a \in \mathbf{Z}_p$. Si $f \in \mathrm{LA}_h$, alors, quel que soit $x_0 \in \mathbf{Z}_p$, on peut développer f sous la forme $f(x) = \sum_{k=0}^{+\infty} a_k(x_0)\left(\frac{x-x_0}{p^h}\right)^k$, où $a_k(x_0)$ est une suite d'éléments de $\mathbf{C}_p$ tendant vers 0 quand k tend vers $+\infty$. Si $f \in \mathrm{LA}_h$ et $x_0 \in \mathbf{Z}_p$, on pose $\|f\|_h = \sup_{k\in\mathbf{N},x_0\in\mathbf{Z}_p}(|a_k(x_0)|)$ ce qui fait de LA_h un espace de Banach p-adique. On note LA l'espace des fonctions localement analytiques sur $\mathbf{Z}_p$ à valeurs dans $\mathbf{C}_p$; c'est la réunion des LA_h car $\mathbf{Z}_p$ est compact.

On dit qu'une fonction sur $\mathbf{Z}_p$ est de classe $\mathscr{C}_u^\infty$ si, pour tout $k \in \mathbf{N}$, la fonction $f^{[k]}$ définie sur $\mathbf{Z}_p \times (\mathbf{Z}_p - \{0\})^k$ par la formule

$$f^{[k]}(x, y_1, \ldots, y_k) = \frac{1}{y_1 \ldots y_k}\Big(\sum_{I\subset\{1,\ldots,k\}} (-1)^{k-|I|} f\Big(x + \sum_{i\in I} y_i\Big)\Big)$$

se prolonge en une fonction continue sur $\mathbf{Z}_p^{k+1}$.

Si f est une fonction de $\mathbf{Z}_p$ dans $\mathbf{C}_p$, soit $a_n(f)$ le n-ième coefficient de Mahler de f ; c'est l'élément de $\mathbf{C}_p$ donné par la formule $a_n(f) = \sum_{i=0}^{n}(-1)^{n-i}\binom{n}{i} f(i)$.

THÉORÈME 1.1 (Mahler). — *f est continue sur $\mathbf{Z}_p$ si et seulement si $a_n(f)$ tend vers 0 quand n tend vers $+\infty$. De plus, la série $\sum_{n=0}^{+\infty} a_n(f)\binom{x}{n}$ tend uniformément vers f sur $\mathbf{Z}_p$ et $\|f\|_{\mathscr{C}^0} = \sup_{n\in\mathbf{N}} |a_n(f)|$.*

Le théorème de Mahler nous donne donc une caractérisation des fonctions continues en termes de ses coefficients de Mahler. Le théorème suivant [1, 4] nous donne une caractérisation analogue des fonctions localement analytiques ou de classe $\mathscr{C}_u^\infty$.

THÉORÈME 1.2. — *(i) $f \in \mathrm{LA}_h$ si et seulement si $v_p(a_n(f)) - v_p([\frac{n}{p^h}]!)$ tend vers $+\infty$ quand n tend vers $+\infty$. De plus, $\|f\|_{\mathrm{LA}_h} = p^{-\inf_n v_p(a_n(f))-v_p([\frac{n}{p^h}]!)}$.*

(ii) $f \in \mathrm{LA}$ si et seulement si $\liminf \frac{1}{n} v_p(a_n(f)) > 0$.

(iii) $f \in \mathscr{C}_u^\infty$ si et seulement si la suite de terme général $(1+n)^k|a_n(f)|$ tend vers 0 quand n tend vers $+\infty$ quel que soit $k \in \mathbf{N}$.

1.1.2. *Distributions continues.* Si A est un espace de Banach p-adique, on appelle distribution continue sur $\mathbf{Z}_p$ à valeurs dans A toute application linéaire de LA dans A dont la restriction à LA_h est continue pour tout $h \in \mathbf{N}$. Les distributions continues sur $\mathbf{Z}_p$ à valeurs dans $\mathbf{C}_p$ forment une algèbre commutative et associative pour le produit de convolution.

On appelle mesure sur $\mathbf{Z}_p$ à valeurs dans A toute application linéaire continue de $\mathscr{C}^0$ dans A. Les mesures sur $\mathbf{Z}_p$ à valeurs dans $\mathbf{C}_p$ forment une sous-algèbre des distributions continues.

A une distribution continue, on associe deux séries formelles

$$\mathscr{A}_\mu(T) = \sum_{n=0}^{+\infty} b_n(\mu) T^n \text{ avec } b_n(\mu) = \int_{\mathbf{Z}_p} \binom{x}{n} \mu(x) \quad \text{et} \quad \mathscr{L}_\mu(t) = \sum_{n=0}^{+\infty} \frac{t^n}{n!} \int_{\mathbf{Z}_p} x^n \mu(x)$$

appelées respectivement transformée d'Amice [2] et transformée de Laplace de μ. On a $\mathscr{L}_\mu(t) = \mathscr{A}_\mu(e^t - 1)$ et

$$\mathscr{A}_\mu(z) = \int_{\mathbf{Z}_p} (1+z)^x \mu(x) \text{ si } |z| < 1 \text{ et } \mathscr{L}_\mu(z) = \int_{\mathbf{Z}_p} e^{zx} \mu(x) \text{ si } |z| < p^{-\frac{1}{p-1}}.$$

Comme les fonctions de la forme ε^x, où ε décrit l'ensemble des racines de l'unité d'ordre une puissance de p, forment une base des fonctions localement constantes sur $\mathbf{Z}_p$, la transformée d'Amice permet de calculer les intégrales du type $\int_{\mathbf{Z}_p} x^k f(x)\,\mu$, où f est une fonction localement constante et $k \in \mathbf{N}$ (mais pas pour $k < 0$). En particulier, on a $\mathbf{1}_{\mathbf{Z}_p^*}(x) = 1 - \frac{1}{p} \sum_{\varepsilon^p=1} \varepsilon^x$, ce qui permet de calculer la transformée d'Amice de la restriction à $\mathbf{Z}_p^*$ d'une distribution μ en termes de celle de μ. On obtient

$$\mathscr{A}_{\mathrm{Res}_{\mathbf{Z}_p^*}(\mu)}(T) = \mathscr{A}_\mu(T) - \frac{1}{p} \sum_{\varepsilon^p=1} \mathscr{A}_\mu(\varepsilon(1+T) - 1).$$

THÉORÈME 1.3. — *(i) L'application qui à une distribution continue μ associe sa transformée d'Amice est un isomorphisme de l'espace des distributions continues sur $\mathbf{Z}_p$ à valeurs dans A sur celui des séries entières à coefficients dans A de rayon de convergence supérieur ou égal à 1.*

(ii) Sa restriction aux mesures est un isomorphisme de l'espace des mesures sur $\mathbf{Z}_p$ à valeurs dans A sur celui des séries entières à coefficients bornés.

DÉFINITION 1.4. — *(i) Si r est un réel positif ou nul, une distribution continue sur $\mathbf{Z}_p$ est dite d'ordre r si la suite de terme général $(1+n)^{-r}|b_n(\mu)|$ est bornée.*

(ii) Une distribution continue μ est dite tempérée si et seulement si il existe $r \in \mathbf{R}_+$ tel que μ soit d'ordre r.

REMARQUE 1.5. — (i) Il résulte du (ii) du théorème précédent qu'une distribution d'ordre 0 n'est rien d'autre qu'une mesure.

(ii) Il résulte du (iii) du théorème 1.2 que l'espace des distributions tempérées est le dual de celui des fonctions de classe $\mathscr{C}_u^\infty$.

(iii) Les distributions tempérées sur $\mathbf{Z}_p$ à valeurs dans $\mathbf{C}_p$ forment une sous-algèbre des distributions continues.

Le théorème 1.3 permet de caractériser les distributions tempérées en termes de leurs transformées d'Amice ce qui permet de construire une distribution tempérée à partir d'une

série entière de rayon de convergence 1 vérifiant des conditions de croissance. La connaissance de la transformée d'Amice d'une distribution est équivalente à la connaissance des $\int_{\mathbf{Z}_p} x^i \mu(x)$ pour $i \in \mathbf{N}$. Le théorème suivant [3, 2, 53, 33, 40] permet de construire une distribution tempérée en ne connaissant que les intégrales du type $\int_{a+p^n\mathbf{Z}_p} x^i \mu(x)$ pour $a \in \mathbf{Z}_p$, $n \in \mathbf{N}$ et $0 \leq i \leq N$. Cette construction est très importante pour les applications arithmétiques.

THÉORÈME 1.6. — *(i) Soit $r \in \mathbf{R}^+$. Une distribution continue est d'ordre r si et seulement si, quel que soit $j \in \mathbf{N}$, la suite de terme général $p^{n(j-r)} \sup_{a\in\mathbf{Z}_p} \| \int_{a+p^n\mathbf{Z}_p} (\frac{x-a}{p^n})^j \mu \|$ est bornée dans $\mathbf{R}$.*

(ii) Réciproquement, soit $N \in \mathbf{N} \cup \{+\infty\}$ strictement supérieur à r et soit μ une application linéaire sur les fonctions localement polynomiales de degré inférieur ou égal à N telle que, quel que soit $j \in [0, N]$, la suite de terme général $p^{n(j-r)} \sup_{a\in\mathbf{Z}_p} \| \int_{a+p^n\mathbf{Z}_p} (\frac{x-a}{p^n})^j \mu \|$ est bornée dans $\mathbf{R}$. Alors il existe une unique distribution d'ordre r coïncidant avec μ sur les fonctions localement polynomiales de degré inférieur ou égal à N.

REMARQUE 1.7. — Pour $r = 0$, le théorème ci-dessus se traduit par le fait qu'une mesure est déterminée par sa valeur sur les fonctions localement constantes, ce qui est une conséquence de la densité de celles-ci dans les fonctions continues.

1.2. Intégration sur $\mathrm{Gal}(\mathbf{Q}(\mu_{p^\infty})/\mathbf{Q})$

On fixe un plongement de $\overline{\mathbf{Q}}$ dans $\mathbf{C}$ et $\mathbf{C}_p$ et on note Frob_∞ la conjugaison complexe. Si $n \in \mathbf{N}$, soit $\varepsilon_n = \exp \frac{2i\pi}{p^n} \in \overline{\mathbf{Q}} \subset \mathbf{C}$. Soit $\mathbf{Q}_\infty = \cup_{n\in\mathbf{N}} \mathbf{Q}(\varepsilon_n)$ l'extension abélienne maximale de $\mathbf{Q}$ non ramifiée en dehors de p et soient $\Gamma = \mathrm{Gal}(\mathbf{Q}_\infty/\mathbf{Q})$ et $\Gamma_n = \mathrm{Gal}(\mathbf{Q}_\infty/\mathbf{Q}(\varepsilon_n))$ si $n \geq 1$. Le caractère cyclotomique χ_{cycl} induit un isomorphisme de Γ sur $\mathbf{Z}_p^*$, ce qui permet de parler de distribution continue ou tempérée ou encore de mesure sur Γ.

On pose $q = 4$ (resp. $q = p$) si $p = 2$ (resp. si $p \neq 2$) et on note Δ le groupe des racines de l'unité contenues dans $\mathbf{Q}_p^*$. Si $x \in \mathbf{Z}_p^*$, on note $\omega(x)$ l'unique élément de Δ vérifiant $x - \omega(x) \in q\mathbf{Z}_p$ et on pose $\langle x \rangle = \omega(x)^{-1}x \in 1 + q\mathbf{Z}_p$. Les applications $x \to \omega(x)$ et $x \to \langle x \rangle$ sont des caractères localement analytiques sur $\mathbf{Z}_p^*$ à valeurs dans $\mathbf{Z}_p^*$.

On sait depuis la thèse de Tate qu'il vaut mieux considérer une fonction L complexe comme une fonction sur les caractères continus du groupe des idèles plutôt que comme fonction d'une variable complexe. De même, une fonction L p-adique est une fonction définie sur les caractères continus $\psi : \Gamma \to \mathbf{C}_p^*$. Un tel caractère est automatiquement localement analytique et ceux de la forme $\eta\langle\chi_{\mathrm{cycl}}\rangle^s$, avec $s \in \mathbf{Z}_p$ et η d'ordre fini, seront particulièrement intéressants pour la suite de l'exposé (en particulier dans le cas $s \in \mathbf{Z}$). De manière idéale, on aimerait bien pouvoir écrire une fonction L p-adique sous la forme $L(\psi) = \int_\Gamma \psi\, \mu$, où μ est une distribution tempérée sur Γ car une fonction définie de cette manière a de bonnes propriétés (elle est par exemple analytique en ψ). Le problème est que les fonctions L p-adiques peuvent avoir des pôles en certains caractères, ce qui nous amène à introduire les notions de pseudo-mesures et pseudo-distributions tempérées.

Les distributions tempérées (resp. les mesures) sur Γ forment une algèbre $\mathscr{D}(\Gamma)$ (resp. $\mathscr{D}_0(\Gamma)$) pour la convolution et si ψ est un caractère de Γ, on a $\int_\Gamma \psi\, \mu\star\mu' = \int_\Gamma \psi\, \mu \cdot \int_\Gamma \psi\, \mu'$, ce qui permet de définir $\int_\Gamma \psi\, \mu$ pour n'importe quel élément μ de l'anneau total des fractions de $\mathscr{D}(\Gamma)$.

On appelle pseudo-mesure sur Γ un élément du localisé de $\mathscr{D}_0(\Gamma)$ en la partie multiplicative engendrée par les images réciproques par χ_{cycl} des mesures μ sur $\mathbf{Z}_p^*$ de la forme $\int_{\mathbf{Z}_p^*} f\mu = f(u) - u^k\eta(u)f(1)$, où k décrit $\mathbf{Z}$, η décrit les caractères d'ordre fini de $\mathbf{Z}_p^*$ et u décrit $1+p\mathbf{Z}_p$. Si μ est une pseudo-mesure sur Γ, la fonction $\psi \to \int_\Gamma \psi$ n'a des pôles qu'en un nombre fini de caractères de Γ de la forme $\eta\chi_{\text{cycl}}^k$, où η est d'ordre fini et $k \in \mathbf{Z}$. L'idée de passer à l'anneau des fractions a été introduite par Serre [50] qui utilise le mot pseudo-mesure dans un sens un peu différent du nôtre.

Si $h \in \mathbf{Z}$, soit ℓ_h l'image réciproque par χ_{cycl} de la distribution $\tilde{\ell}_h$ sur $\mathbf{Z}_p^*$ définie par $\int_{\mathbf{Z}_p^*} f\,\tilde{\ell}_h = f'(1) + hf(1)$. On appelle pseudo-distribution tempérée sur Γ un élément du localisé de $\mathscr{D}(\Gamma)$ en la partie multiplicative engendrée par les ℓ_h. Un petit calcul nous donne $\int_\Gamma \eta\langle\chi_{\text{cycl}}\rangle^s \ell_h = s+h$ si η est un caractère d'ordre fini de Γ et $s \in \mathbf{Z}_p$; on en déduit le fait que si μ est une pseudo-distribution tempérée, alors $\int_\Gamma \eta\langle\chi_{\text{cycl}}\rangle^s\,\mu = \frac{F(s)}{P(s)}$, où F est une fonction analytique sur $\mathbf{Z}_p$ et P est un polynôme dont toutes les racines sont dans $\mathbf{Z}$. L'anneau des pseudo-distributions tempérées sera noté $\tilde{\mathscr{D}}(\Gamma)$.

Le groupe Γ étant le produit direct du groupe fini Δ par un groupe isomorphe à $\mathbf{Z}_p$, l'algèbre $\tilde{\mathscr{D}}(\Gamma)$ se décompose en un produit d'algèbres correspondant aux caractères du groupe Δ. Nous aurons en particulier besoin des idempotents

$$e^+ = \frac{1}{2}(\delta_1 + \delta_{\text{Frob}_\infty}) \quad \text{et} \quad e^- = \frac{1}{2}(\delta_1 - \delta_{\text{Frob}_\infty})$$

qui permettent de décomposer toute pseudo-mesure μ sous la forme $\mu = \mu^+ + \mu^-$, avec $\mu^+ = e^+\mu$ et $\mu^- = e^-\mu$.

1.3. **Fonctions L p-adiques attachées aux caractères de Dirichlet**

On verra indifféremment un caractère de Dirichlet de conducteur d comme un caractère de $(\mathbf{Z}/d\mathbf{Z})^*$ ou comme un caractère de $\mathscr{G}_{\mathbf{Q}}$ se factorisant à travers $\text{Gal}(\mathbf{Q}(\mu_d)/\mathbf{Q})$. Si χ_1 et χ_2 sont deux tels caractères, on note $\chi_1 \otimes \chi_2$ le caractère de Dirichlet primitif attaché à leur produit (c'est juste leur produit si on utilise la seconde interprétation). Si χ est un caractère de Dirichlet de conducteur d, soit $G(\chi) = \sum_{x\in(\mathbf{Z}/d\mathbf{Z})^*} \chi(x)e^{\frac{2i\pi x}{d}}$ la somme de Gauss associée à χ. Soit finalement $L(\chi,s) = \sum_{n=1}^{+\infty} \frac{\chi(n)}{n^s}$ la fonction L complexe attachée à χ. Cette fonction admet un prolongement méromorphe à tout le plan complexe, holomorphe en dehors d'un pôle simple en $s=1$ si χ est le caractère trivial et on sait que $L(\chi,k)$ est un nombre algébrique si $k \leq 0$ et que $(2i\pi)^{-k}L(\chi,k)$ est un nombre algébrique si $\chi(\text{Frob}_\infty)(-1)^k = 1$ et $k > 0$. Finalement, en $s=1$, on a la formule suivante :

$$\text{(1)} \qquad \text{si } d \neq 1, \text{ alors } \quad L(\chi,1) = -\frac{1}{G(\chi^{-1})} \sum_{x \bmod d} \chi^{-1}(x)\log(1-e^{\frac{2i\pi x}{d}}).$$

On peut interpoler p-adiquement les valeurs de la fonction L attachée à χ pour construire sa fonction L p-adique. De manière précise, on a le résultat suivant.

THÉORÈME 1.8. — *Soit χ un caractère de Dirichlet de conducteur d premier à p. Il existe une unique pseudo-mesure μ_χ sur Γ qui est une mesure si $\chi \neq 1$, telle que l'on ait*

$$\int_\Gamma \chi_{\mathrm{cycl}}^k \mu_\chi = \begin{cases} (1-\chi(p)p^{-k})L(\chi,k) & \text{si } k \leq 0 \text{ et } \chi(\mathrm{Frob}_\infty)(-1)^k = -1, \\ \frac{\Gamma(k)}{(2i\pi)^k}(1-\chi^{-1}(p)p^{k-1})L(\chi,k) & \text{si } k \geq 1 \text{ et } \chi(\mathrm{Frob}_\infty)(-1)^k = 1. \end{cases}$$

De plus la fonction L p-adique $\psi \to L_p(\chi\otimes\psi) = \int_\Gamma \psi\,\mu_\chi$ vérifie les propriétés suivantes :

(i) elle est analytique sauf si χ est le caractère trivial auquel cas elle n'est analytique qu'en dehors des caractères 1 et χ_{cycl} en lesquels elle a des pôles simples de résidus respectifs $-\frac{1}{2}(1-\frac{1}{p})$ et $1-\frac{1}{p}$,

(ii) elle vérifie l'équation fonctionnelle

$$L_p(\chi\otimes\psi) = 2^{-\chi\otimes\psi(\mathrm{Frob}_\infty)}G(\chi)\psi(\sigma_d)L_p(\chi^{-1}\otimes\psi^{-1}\chi_{\mathrm{cycl}}),$$

où σ_d est l'élément de Γ tel que $\chi_{\mathrm{cycl}}(\sigma_d) = d$,

(iii) si $m \geq 1$ et si η est un caractère d'ordre fini de Γ trivial modulo Γ_m mais pas modulo Γ_{m-1}, alors

$$L_p(\chi\otimes\eta\chi_{\mathrm{cycl}}^k) = \begin{cases} L(\eta,k) & \text{si } k \leq 0 \text{ et } \eta\otimes\chi(\mathrm{Frob}_\infty)(-1)^k = -1, \\ \frac{\Gamma(k)}{(2i\pi)^k}\frac{p^{mk}}{G(\eta)}L(\eta,k) & \text{si } k \geq 1 \text{ et } \eta\otimes\chi(\mathrm{Frob}_\infty)(-1)^k = 1. \end{cases}$$

REMARQUE 1.9. — (i) L'existence de pôles simples en 1 et χ_{cycl} se traduit par les formules $\lim_{s\to 0} sL_p(\langle\chi_{\mathrm{cycl}}\rangle^s) = -\frac{1}{2}(1-\frac{1}{p})$ et $\lim_{s\to 0} sL_p(\chi_{\mathrm{cycl}}\langle\chi_{\mathrm{cycl}}\rangle^s) = 1-\frac{1}{p}$.

(ii) Si $i \in \mathbf{Z}/(p-1)\mathbf{Z}$ est impair (resp. $i = 1$ si $p = 2$), soit $\zeta_{p,i}$ la fonction définie sur $\mathbf{Z}_p$ par la formule $\zeta_{p,i}(s) = L_p(\omega^i\langle\chi_{\mathrm{cycl}}\rangle^s)$. Il résulte du théorème précédent que cette fonction est analytique sur $\mathbf{Z}_p$ (resp. méromorphe sur $\mathbf{Z}_p$, holomorphe en dehors de $s = 1$ où elle a un pôle simple de résidu $1-\frac{1}{p}$) si $i \neq 1$ (resp. $i = 1$) et que l'on a $\zeta_{p,i}(-n) = (1-p^n)\zeta(-n)$ si $n \in \mathbf{N}$ vérifie $-n \equiv i$ modulo $p-1$. La fonction $\zeta_{p,i}$ est la i-ème branche de la fonction zêta de Kubota-Leopoldt.

(iii) L'équation fonctionnelle de la fonction L p-adique nous fournit la formule

$$L(\eta,k) = \frac{1}{2}\frac{\Gamma(1-k)}{(2i\pi)^{1-k}}\frac{p^{m(1-k)}}{G(\eta^{-1})}L(\eta^{-1},1-k),$$

si $k \geq 0$ et $\eta(\mathrm{Frob}_\infty)(-1)^k = -1$. Cette formule est équivalente à l'équation fonctionnelle de la fonction L complexe aux entiers.

(iv) On voit que pour obtenir une propriété de continuité p-adique, on a été forcé de modifier un peu les valeurs des fonctions L. En particulier, on a dû rajouter ou non selon les cas, un "facteur à l'infini" $\frac{\Gamma(k)}{(2i\pi)^k}$, un facteur d'Euler en p et des constantes locales (sommes de Gauss). Le facteur d'Euler en p est soit celui du point que l'on considère soit

celui du point qui lui correspond via l'équation fonctionnelle ; il trouvera une explication conceptuelle au chapitre suivant.

(v) Il est assez remarquable que la fonction L p-adique continue à être reliée à la fonction L complexe même aux caractères dont on ne s'est pas servi pour construire la fonction L p-adique (cf. (iii) du théorème ci-dessus). Le résultat suivant que l'on pourra comparer avec la formule (1) est une illustration du même phénomène en un point où l'on n'a plus de résultat d'algébricité.

PROPOSITION 1.10 (Leopoldt). — *Si* $\chi(\mathrm{Frob}_\infty)=1$, *alors*

$$L_p(\chi\otimes\chi_{\mathrm{cycl}})=\frac{-1}{G(\chi^{-1})}(1-\frac{\chi(p)}{p})\sum_{x \bmod d}\chi^{-1}(x)\log_p(e^{\frac{2i\pi x}{d}}-1).$$

Plus généralement, si $m\geq 1$ *et* η *est un caractère de* Γ *trivial modulo* Γ_m *mais pas* Γ_{m-1} *et tel que* $\chi\otimes\eta(\mathrm{Frob}_\infty)=1$, *alors*

$$L_p(\chi\otimes\eta\chi_{\mathrm{cycl}}))=\frac{-1}{G((\chi\otimes\psi)^{-1})}\sum_{x \bmod dp^m}(\chi\otimes\psi)^{-1}(x)\log_p(e^{\frac{2i\pi x}{dp^m}}-1).$$

Il y a un nombre considérable de manières de démontrer ces résultats. Considérons le cas du caractère trivial (le cas général se traite exactement de la même manière) et notons $\mu_{\mathrm{K-L}}=\mu^+_{\mathrm{K-L}}+\mu^-_{\mathrm{K-L}}$ la pseudo-mesure correspondante. La manière la plus efficace pour construire $\mu^-_{\mathrm{K-L}}$, compte tenu de la caractérisation des mesures par leur transformée d'Amice consiste à montrer que si $a\in\mathbf{Z}_p$, alors il existe une mesure sur $\mathbf{Z}_p$ dont la transformée de Laplace est $\frac{1}{e^t-1}-\frac{a}{e^{at}-1}$. Cette mesure μ_a vérifie alors les formules $\int_{\mathbf{Z}_p}x^n\mu_a=(-1)^n(1-a^{1+n})\zeta(-n)$ si $n\leq 0$ comme on le voit à partir de la représentation intégrale $(1-a^{1-s})\zeta(s)=\frac{1}{\Gamma(s)}\int_0^{+\infty}\Big(\frac{1}{e^t-1}-\frac{a}{e^{at}-1}\Big)t^s\frac{dt}{t}$. Pour obtenir une mesure sur Γ à partir de μ_a, on commence par restreindre μ_a à $\mathbf{Z}_p^*$ (un petit calcul montre que cela fait sortir le facteur d'Euler en p) puis on se débarasse de la dépendance en a (et du facteur parasite $1-a^{1+n}$) en divisant par $\delta_a-\delta_1$ et on obtient une pseudo-distribution que l'on tire sur Γ grâce à $\chi_{\mathrm{cycl}}^{-1}$ et qui n'est autre que $\mu^-_{\mathrm{K-L}}$. Pour construire $\mu^+_{\mathrm{K-L}}$, on utilise les entiers positifs en partant de la formule $-\sum_{k\in\mathbf{Z}}\frac{1}{t-2i\pi k}=\frac{1}{2}\mathrm{coth}\frac{t}{2}$, d'où l'on tire $(1-a^{-2n})\frac{\Gamma(2n)}{(2i\pi)^{2n}}\zeta(2n)=\frac{1}{2}\big(\frac{d}{dt}\big)^{2n-1}\Big(\frac{1}{2}\mathrm{coth}\frac{t}{2}-\frac{1}{2a}\mathrm{coth}\frac{t}{2a}\Big)_{|t=0}$.

La construction la plus pertinente pour cet exposé est celle de Coleman [13] qui repose sur les "séries de Coleman" dont la construction est rappelée dans la proposition suivante.

PROPOSITION 1.11. — *Si* $u=(u_n)_{n\in\mathbf{N}}$ *est un élément de la limite projective* $\varprojlim\mathcal{O}^*_{\mathbf{Q}_p(\varepsilon_n)}$ *des* $\mathcal{O}^*_{\mathbf{Q}_p(\varepsilon_n)}$ *relativement aux applications normes, il existe une unique série* $\mathrm{Col}_u(T)$ *élément de* $(\mathbf{Z}_p[[T]])^*$ *telle que l'on ait* $\mathrm{Col}_u(\varepsilon_n-1)=u_n$ *quel que soit* $n\geq 1$.

Pour retrouver la fonction zêta de Kubota-Leopoldt à partir de ce résultat, commençons par remarquer que les unités cyclotomiques fournissent des éléments de $\varprojlim\mathcal{O}^*_{\mathbf{Q}(\varepsilon_n)}$ et donc

de $\varprojlim \mathcal{O}^*_{\mathbf{Q}_p(\varepsilon_n)}$. De manière précise, si $\gamma \in \Gamma$, la collection $u_\gamma = (u_{\gamma,n})_{n\in\mathbf{N}}$, où $u_{\gamma,n} = \frac{\gamma(\varepsilon_n)-1}{\varepsilon_n-1}$ si $n \geq 1$, est un élément de $\varprojlim \mathcal{O}^*_{\mathbf{Q}(\varepsilon_n)}$ et si $a = \chi_{\mathrm{cycl}}(\gamma)$, alors $\mathrm{Col}_{u_\gamma}(T) = \frac{(1+T)^a-1}{T}$.

D'autre part, si $u \in \varprojlim \mathcal{O}^*_{\mathbf{Q}_p(\varepsilon_n)}$, Il existe une unique distribution tempérée λ_u sur $\mathbf{Z}_p$ dont la transformée d'Amice est $\log(\mathrm{Col}_u(T))$ et une comparaison des transformées d'Amice montre que si μ_a est la mesure définie ci-dessus, alors $\mu_a = x\lambda_{u_\gamma}$, ce qui nous fournit d'une part une construction de la fonction zêta de Kubota-Leopoldt et d'autre part une manière de calculer $\int_{\mathbf{Z}_p} f(x)x^{-1}\mu_a$ si f est une fonction localement constante ; on en tire la proposition 1.10.

1.4. **Fonctions L p-adiques attachées aux formes modulaires**

Soit $f(z) = \sum_{n=1}^{+\infty} a_n e^{2i\pi nz}$ une forme modulaire de poids $2k$ pour $\mathrm{SL}_2(\mathbf{Z})$. On suppose que f est propre pour les opérateurs de Hecke et $a_1 = 1$. Le sous-corps de $\mathbf{C}$ engendré par les a_n est un corps de nombres totalement réel que nous noterons $\mathbf{Q}[f]$.

Si χ est un caractère de Dirichlet de conducteur m, on pose

$$L(f\otimes\chi, s) = \sum_{n=1}^{+\infty} \frac{\chi(n)a_n}{n^s} = \prod_\ell \frac{1}{1 - a_\ell(\chi(\ell)\ell^{-s}) + \ell^{2k-1}(\chi(\ell)\ell^{-s})^2}.$$

La fonction $L(f\otimes\chi, s)$ a un prolongement analytique à $\mathbf{C}$ tout entier et vérifie l'équation fonctionnelle

$$m^s \frac{\Gamma(s)}{(2\pi)^s}\frac{L(f,\chi,s)}{G(\chi)} = (-1)^k\chi(-1)m^{2k-s}\frac{\Gamma(2k-s)}{(2\pi)^{2k-s}}\frac{L(f,\chi^{-1},2k-s)}{G(\chi^{-1})}.$$

On peut fabriquer [32, 3, 53] une fonction L p-adique à partir des nombres

$$\Lambda(f\otimes\chi, j) = m^j\frac{\Gamma(j)}{(-2i\pi)^j}L(f\otimes\chi, j),$$

pour $j \in \{1,\ldots,2k-1\}$ et χ variant parmi les caractères de Dirichlet de conducteur une puissance de p. Le principal ingrédient de la construction est le théorème suivant.

THÉORÈME 1.12 (Manin). — *Soit f une forme modulaire cuspidale de poids $2k$ pour $\mathrm{SL}_2(\mathbf{Z})$ propre sous l'action des opérateurs de Hecke. Il existe alors $\Omega_f^- \in \mathbf{R}$ et $\Omega_f^+ \in i\mathbf{R}$ tels que l'on ait*

$$\frac{\Lambda(f\otimes\chi, j)}{G(\chi)} \in \begin{cases} \Omega_f^-\mathbf{Q}[f,\chi] & \text{si } j \in \{1,\ldots,2k-1\} \text{ et } \chi(-1)(-1)^{j-1} = 1, \\ \Omega_f^+\mathbf{Q}[f,\chi] & \text{si } j \in \{1,\ldots,2k-1\} \text{ et } \chi(-1)(-1)^{j-1} = -1 \end{cases}.$$

La démonstration de ce théorème s'appuie sur l'étude des intégrales (appelées symboles modulaires) du type $\int_r^{i\infty} f(z)P(z)dz$, où $r \in \mathbf{Q}$ et P est un polynôme de degré inférieur ou égal à $2k-2$ à coefficients entiers. On montre que, si $a \in \mathbf{Q}^*$ et $b \in \mathbf{Q}$, alors $\int_0^{i\infty} f(az+b)z^j$ a une partie réelle appartenant à $\Omega_f^-\overline{\mathbf{Q}}$ et une partie imaginaire appartenant à $\Omega_f^+\overline{\mathbf{Q}}$.

CONVENTION 1. — si $a \in \mathbf{Z} - \{0\}$ et $b \in \mathbf{Z}$, on note $\int_0^{i\infty} f(az+b)z^j$ l'élément de $\mathbf{C}_p$ défini en faisant la somme de la partie réelle de $\int_0^{i\infty} f(az+b)z^j$ divisée par Ω_f^- et de sa partie imaginaire divisée par Ω_f^+.

THÉORÈME 1.13. — *Soit f une forme modulaire de poids $2k$ pour $\mathrm{SL}_2(\mathbf{Z})$ propre pour les opérateurs de Hecke et normalisée. Soient a_p la valeur propre de T_p, α une racine du polynôme $X^2 - a_p X + p^{2k-1} = 0$ et $\beta = \frac{p^{2k-1}}{\alpha}$ l'autre racine. Si $v_p(\alpha) < 2k-1$, il existe une unique distribution $\tilde{\mu}_{f,\alpha}$ d'ordre $v_p(\alpha)$ telle que si $n \in \mathbf{N}$, $a \in \mathbf{Z}_p$ et $i \in \{0, \dots, 2k-2\}$, alors*

$$\int_{a+p^n\mathbf{Z}_p} x^i \tilde{\mu}_{f,\alpha}(x) = \frac{1}{(p\alpha)^n} \int_0^{i\infty} \Big(f\big(\frac{z-a}{p^n}\big) - \frac{p^{2k-1}}{\alpha} f\big(\frac{z-a}{p^{n-1}}\big)\Big) z^i dz.$$

REMARQUE 1.14. — La distribution $\tilde{\mu}_{f,\alpha}$ est vraiment d'ordre $v_p(\alpha)$; en particulier, si $v_p(\alpha) \neq 0$, ce n'est pas une mesure. Dans le cas où $v_p(\alpha) = 2k-1$, on n'a pas assez de valeurs pour garantir l'unicité d'une distribution tempérée d'ordre $v_p(\alpha)$ vérifiant les formules de la proposition, ce qui rend le résultat inutilisable tel quel. D'autre part, a_p étant un entier, il y a deux cas possibles. Soit $v_p(a_p) = 0$ (cas ordinaire) et une des racines de $X^2 - a_p X + p^{2k-1}$ est de valuation 0 tandis que l'autre est de valuation $2k-1$, soit $v_p(a_p) > 0$ (cas supersingulier) et les deux racines de $X^2 - a_p X + p^{2k-1}$ sont de valuation strictement comprise entre 0 et $2k-1$.

On note $\mu_{f,\alpha}$ la distribution tempérée sur Γ image réciproque de la restriction de $\tilde{\mu}_{f,\alpha}$ à $\mathbf{Z}_p^*$ et on définit la fonction L p-adique de f (associée à α) par la formule $L_p(f \otimes \psi, \alpha) = \int_\Gamma \psi(x)\mu_{f,\alpha}(x)$.

PROPOSITION 1.15. — *La fonction L p-adique est reliée à la fonction L complexe par les formules suivantes*

(i) Si $i \in \{0, \dots, 2k-2\}$,

$$L_p(f \otimes \chi_{\mathrm{cycl}}^i, \alpha) = (1 - \beta p^{-(i+1)})(1 - \beta p^{-(2k-i-1)})\Lambda(f, i+1).$$

(ii) Si $m \geq 1$ et η est un caractère d'ordre fini de Γ trivial modulo Γ_m mais pas modulo Γ_{m-1} et si $i \in \{0, \dots, 2k-2\}$, alors

$$L_p(f \otimes \eta\chi_{\mathrm{cycl}}^i, \alpha) = \frac{1}{\alpha^m} \frac{\Lambda(f, \eta^{-1}, i+1)}{G(\eta^{-1})}.$$

(iii) De plus, la fonction L p-adique attachée à f vérifie l'équation fonctionnelle

$$L_p(f \otimes \psi, \alpha) = -\psi(-1) L_p(f \otimes \chi_{\mathrm{cycl}}^{2k-2}\psi^{-1}, \alpha)$$

REMARQUE 1.16. — i) Le facteur d'Euler en p de $L(f,s)$ est

$$\frac{1}{1 - a_p p^{-s} + p^{2k-1}p^{-2s}} = \Big(\frac{1}{1-\alpha p^{-s}}\Big)\Big(\frac{1}{1-\beta p^{-s}}\Big).$$

On voit donc que le facteur d'Euler que l'on doit rajouter pour rendre les choses p-adiquement interpolables est le produit d'un morceau du facteur d'Euler en $i+1$ et d'un

morceau de celui en $2k-i-1$ qui est le point correspondant à $i+1$ via l'équation fonctionnelle.

(ii) La théorie effleurée ci-dessus s'étend aux formes de niveau quelconque ([33] par exemple) et l'équation fonctionnelle est alors un peu plus compliquée.

2. LA FONCTION L p-ADIQUE D'UN MOTIF

2.1. Déterminants et discriminants

Rappelons que si K est un corps, on peut associer à un K-espace vectoriel V de dimension finie $d(V)$ son déterminant $\det_K V$ qui est une K-droite (i.e. un K-espace vectoriel de dimension 1).

Si ω est une famille de $d(V)$ vecteurs de V, on note $[\omega]$ son déterminant ; c'est un élément de $\det_K V$. Les K-droites forment un groupe sous le produit tensoriel dont l'élément neutre est K et l'inverse d'une droite D est la droite duale D^*. Si ω est une base d'un K-espace vectoriel de dimension finie V et ω^* est la base de V^* duale de ω, alors $[\omega][\omega^*]=1$. Si $0\to V_1\to V\to V_2\to 0$ est une suite exacte de K-espaces vectoriels de dimensions finies, alors $\det_K V$ s'identifie canoniquement (au signe près) à $(\det_K V_1)\otimes(\det_K V_2)$ (si ω, ω_1 et ω_2 sont respectivement des bases sur K de V, V_1 et V_2 et $\widehat{\omega}_2$ un relèvement de ω_2 dans V, le déterminant de la famille $(\omega_1,\widehat{\omega}_2)$ dans la base ω ne dépend pas du choix de $\widehat{\omega}_2$).

Soit A un anneau principal contenu dans K. Si T est un A-module de type fini et de torsion, notons $|T|$ le produit des diviseurs élémentaires de T ; c'est un élément de A bien défini à multiplication près par une unité de A. Si $A=\mathbf{Z}$, alors $|T|$ est le cardinal de T. Si T est un A-module de type fini et $V=K\otimes A$, alors $\det_A T$ est un sous-A-module de $\det_K V$ dont on note $[T]$ un générateur. Si T est de torsion, alors $V=0$, $\det_K V=K$ et $[T]=|T|^{-1}$. Dans le cas général, si ω est une famille d'éléments de T formant une base de V, alors $[T]\sim|T/\omega|^{-1}[\omega]$ (de manière générale, on note de la même manière une famille d'éléments d'un A-module et le sous-A-module qu'elle engendre).

Si $0\to V_1\to V_2\to\cdots\to V_n\to 0$ est une suite exacte de K-espaces vectoriels, $\otimes_{i=1}^n(\det_K V_i)^{(-1)^i}$ s'identifie canoniquement (au signe près) à K et si T_i est un A-module de type fini tel que l'on ait $K\otimes T_i=V_i$ pour $1\le i\le n$, la quantité $\prod_{i=1}^n[T_i]^{(-1)^i}\in K^*$ est un élément de K bien défini à multiplication près par une unité de A. Il faut quand même faire attention au fait que les flèches qui interviennent dans la suite exacte ont une grande influence sur le résultat, ce qui n'apparaît pas sur la notation. Si $n=2$, la quantité $[T_1][T_2]^{-1}$ sera parfois notée $[T_2:T_1]$ et appelée l'indice généralisé de T_1 dans T_2 (si $A=\mathbf{Z}$ et l'application de V_1 dans V_2 est induite par une application injective de T_1 dans T_2, on retombe bien sur l'indice de (l'image de) T_1 dans T_2).

Un accouplement parfait $\langle\ ,\ \rangle$ entre deux K-espaces vectoriels V_1 et V_2 de (même) dimension finie induit un isomorphisme de $(\det_K V_1)\otimes(\det_K V_2)$ sur K. Si T_1 et T_2 sont

deux A-modules de type fini tels que l'on ait $V_i = K \otimes T_i$ pour $i = 1, 2$, on pose $\langle T_1, T_2 \rangle = [T_1] \otimes [T_2]$; c'est un élément de K bien défini à multiplication près par une unité de A et si ω_i est une famille d'éléments de T_i formant une base de V_i, alors $\langle T_1, T_2 \rangle$ est aussi le discriminant de $\langle\ ,\ \rangle$ dans les bases ω_1 et ω_2 multiplié par $|T_1/\omega_1|^{-1}|T_2/\omega_2|^{-1}$.

Finalement, si T est un A-module, on note $T^\vee = \mathrm{Hom}(T, \mathrm{Fr}(A)/A)$ le dual de Pontryagin de T et $T^* = \mathrm{Hom}(T, A)$.

2.2. **Motifs**

Si K est un corps, un motif (resp. un motif pur) défini sur K est un morceau de la cohomologie d'une variété algébrique (resp. variété projective lisse) définie sur K. Nous ne nous intéresserons qu'au cas où K est un corps de nombres et même, pour l'énoncé des conjectures, qu'au cas où $K = \mathbf{Q}$, le cas d'un corps de nombres quelconque pouvant s'y ramener par restriction des scalaires. On associe à un motif défini sur $\mathbf{Q}$ deux $\mathbf{Q}$-espaces vectoriels M_B et M_{dR} de même dimension $d(M)$ appelés respectivement réalisation de Betti et réalisation de de Rham de M et l'entier $d(M)$ est appelé "dimension de M". Ces deux espaces vectoriels sont munis de structures supplémentaires et d'isomorphismes de comparaison dont les propriétés sont rappelées ci-dessous (il faut faire attention au fait que certaines des propriétés énoncées ci-dessous ne sont en fait que partiellement connues).

(Mot 1) M_{dR} est muni d'une filtration décroissante par des sous-$\mathbf{Q}$-espaces vectoriels $(M_{\mathrm{dR}}^i)_{i \in \mathbf{Z}}$ et on appelle espace tangent à M le $\mathbf{Q}$-espace vectoriel $t_M = M_{\mathrm{dR}}/M_{\mathrm{dR}}^0$.

(Mot 2) M_B est muni d'une action de $\mathscr{G}_{\mathbf{R}} = \mathrm{Gal}(\mathbf{C}/\mathbf{R})$ et a donc une décomposition sous la forme $M_B = M_B^+ \oplus M_B^-$, où M_B^+ (resp. M_B^-) est le sous-espace propre de la conjugaison complexe associée à la valeur propre 1 (resp. -1). On note $d^+(M)$ et $d^-(M)$ les dimensions respectives de M_B^+ et M_B^-.

(Mot 3) On dispose d'un isomorphisme "de périodes complexes" $\iota_{M,\infty} : (\mathbf{C} \otimes M_B)^{\mathscr{G}_{\mathbf{R}}} \to \mathbf{R} \otimes M_{\mathrm{dR}}$ qui, composé avec l'injection de $\mathbf{R} \otimes M_B^+$ dans $(\mathbf{C} \otimes M_B)^{\mathscr{G}_{\mathbf{R}}}$ et la surjection de $\mathbf{R} \otimes M_{\mathrm{dR}}$ sur $\mathbf{R} \otimes t_M$, nous fournit une application $\alpha_M : \mathbf{R} \otimes M_B^+ \to \mathbf{R} \otimes t_M$ et une suite exacte dite "suite exacte tautologique"

$$0 \to \ker \alpha_M \to \mathbf{R} \otimes M_B^+ \to \mathbf{R} \otimes t_M \to \mathrm{coker} \alpha_M \to 0.$$

(Mot 4) Si p est un nombre premier, $M_{B,p} = \mathbf{Q}_p \otimes M_B$ est muni d'une action de $\mathscr{G}_{\mathbf{Q}}$ dont la restriction à $\mathscr{G}_{\mathbf{R}}$ est celle que l'on avait sur M_B. Les $M_{B,p}$ forment un système compatible de représentations, c'est-à-dire que si ℓ est un nombre premier, si I_ℓ désigne le sous-groupe d'inertie de $\mathscr{G}_{\mathbf{Q}_\ell}$ et $\mathrm{Frob}_\ell \in \mathscr{G}_{\mathbf{Q}_\ell}$ est un élément de Frobenius, alors le déterminant de $1 - X\mathrm{Frob}_\ell^{-1}$ agissant sur $(M_{B,p})^{I_\ell}$ est un polynôme à coefficients dans $\mathbf{Q}$ indépendant de $p \neq \ell$; nous le noterons $E_\ell(M, X)$.

(Mot 5) Le motif M a une fonction L définie par le produit eulérien $L(M, s) = \prod_\ell E_\ell(M, \ell^{-s})^{-1}$ qui converge pour $\mathrm{Re}(s) \gg 0$.

On conjecture que $L(M, s)$ a un prolongement méromorphe à tout le plan complexe et les conjectures de Deligne, Beilinson, Scholl, Bloch et Kato, Fontaine et Perrin-Riou

rappelées ci-dessous prédisent le comportement de $L(M,s)$ en $s=0$. Nous avons choisi une présentation proche de celle de Fontaine et Perrin-Riou [21, 19]; une référence agréable pour l'historique de la conjecture de Beilinson est [37].

Si $n \in \mathbf{Z}$, on note $M(n)$ le tordu n fois (à la Tate) de M. En tant que $\mathbf{Q}$-espace vectoriel, on a $M(n)_{\mathrm{dR}} = M_{\mathrm{dR}}$, mais la filtration de Hodge est décalée : $M(n)^i_{\mathrm{dR}} = M^{n+i}_{\mathrm{dR}}$ si $i \in \mathbf{Z}$. On a aussi $M(n)_B = M_B$ en tant que $\mathbf{Q}$-espace vectoriel mais l'action de la conjugaison complexe est multipliée par $(-1)^n$ et $\iota_{M(n),\infty} = (2i\pi)^n \iota_{M,\infty}$; en particulier $M(n)^+_B = M_B^{(-1)^n}$. Finalement, l'action de $\mathscr{G}_{\mathbf{Q}}$ sur $M(n)_{B,p}$ est obtenue en multipliant celle sur $M_{B,p}$ par χ^n_{cycl} et la fonction L de $M(n)$ est reliée à celle de M par la formule $L(M(n),s) = L(M,s+n)$.

Soit M^* le dual de M. Les fonctions L de M et $M^*(1)$ sont conjecturalement reliées par une équation fonctionnelle du type $L(M,s) = *L(M^*(1),-s)$, où $*$ est un quotient de produit de fonctions Γ et d'exponentielles que l'on peut décrire explicitement en faisant intervenir la filtration de Hodge et certaines "constantes locales" généralisant les sommes de Gauss.

On note $H^0(M)$ le plus grand sous-motif de M somme directe de copies du motif trivial $\mathbf{Q}(0)$ et $H^1(M)$ le groupe des extensions[1] de $\mathbf{Q}(0)$ par M. Si $x \in H^1(M)$, soit $0 \to M \to M_x \to \mathbf{Q}(0) \to 0$ l'extension correspondante. Il lui correspond, pour chaque nombre premier p, une extension $0 \to M_{B,p} \to (M_x)_{B,p} \to \mathbf{Q}_p \to 0$ de représentations p-adiques. Si $e \in (M_x)_{B,p}$ est un relèvement de $1 \in \mathbf{Q}_p$, le cocycle $\sigma \to (\sigma - 1)e$ sur $\mathscr{G}_{\mathbf{Q}}$ est à valeurs dans $M_{B,p}$ et la classe de ce cocycle dans $H^1(\mathscr{G}_{\mathbf{Q}}, M_{B,p})$ ne dépend pas des choix que l'on a faits; elle sera notée $x_{B,p}$. Si ℓ est un nombre premier, on dit que x a bonne réduction en ℓ si, pour tout $p \neq \ell$, la restriction de $x_{B,p}$ au sous-groupe d'inertie I_ℓ de $\mathscr{G}_{\mathbf{Q}_\ell} \subset \mathscr{G}_{\mathbf{Q}}$ est triviale. Si S est un ensemble fini de nombres premiers, on note $H^1_{f,S}$ le sous-groupe de $H^1(M)$ des extensions de $\mathbf{Q}(0)$ par M ayant bonne réduction en dehors de S et on note $H^1_f(M)$ le sous-groupe de $H^1(M)$ correspondant à $S = \emptyset$.

Si $x \in H^1(M)$, soit $e_{\mathrm{dR}} \in (M_x)^0_{\mathrm{dR}}$ relevant $1 \in \mathbf{Q}(0)^0_{\mathrm{dR}} = \mathbf{Q}$ et $e_B \in (M_x)^+_B$ relevant $1 \in \mathbf{Q}(0)^+_B = \mathbf{Q}$. L'élément $e_{\mathrm{dR}} - \iota_{M_x,\infty}(e_B)$ de $\mathbf{R} \otimes (M_x)_{\mathrm{dR}}$ appartient à $\mathbf{R} \otimes M_{\mathrm{dR}}$ et son image $x_{\mathrm{dR},\infty}$ dans $\mathrm{coker}\alpha_M$ ne dépend d'aucun des choix que l'on a fait; on a donc défini de cette manière une application $v_M : \mathbf{R} \otimes H^1(M) \to \mathrm{coker}\,\alpha_M$.

Maintenant, dans la dualité $M_{\mathrm{dR}} \times M^*(1)_{\mathrm{dR}} \to \mathbf{Q}(1)_{\mathrm{dR}} = \mathbf{Q}$, les espaces M^0_{dR} et $M^*(1)^0_{\mathrm{dR}}$ sont orthogonaux, le dual de t_M s'identifie à $M^*(1)^0_{\mathrm{dR}}$, celui de $t_{M^*(1)}$ à M^0_{dR} et, après tensorisation par $\mathbf{R}$, le dual de $\mathrm{coker}\,\alpha_{M^*(1)}$ s'identifie à $\ker\alpha_M$. On en déduit donc par dualité une application naturelle $u^*_{M^*(1)} : \ker\alpha_M \to \mathbf{R} \otimes H^1(M)^*$.

[1] Nous avons choisi d'utiliser le point de vue des extensions de motifs plutôt que de donner une définition précise en termes de K-théorie, de groupes de Chow supérieurs ou de cohomologie motivique car la plupart des flèches dont nous aurons besoin ont une description plus agréable. Le prix à payer est qu'il est impossible de démontrer quoi que ce soit : par exemple, on ne sait pas démontrer qu'une extension de motifs de $\mathbf{Q}(0)$ par $\mathbf{Q}(1)$ est de type Kummer alors que l'énoncé correspondant en K-théorie est un théorème (facile).

Finalement, on définit via la théorie des biextensions, un accouplement "hauteur" $H^1_f(M) \times H^1_f(M^*(1)) \to \mathbf{R}$ et on conjecture que cet accouplement donne naissance à la suite exacte dite "suite exacte fondamentale de M"

$$0 \to H^0(M)_{\mathbf{R}} \to \ker \alpha_M \to H^1_f(M^*(1))^*_{\mathbf{R}} \to H^1_f(M)_{\mathbf{R}} \to \operatorname{coker} \alpha_M \to H^0(M^*(1))^*_{\mathbf{R}} \to 0.$$

Soient ω_{tg} une base de t_M, ω_B^+ une base de M_B^+, $\omega_{f,M}$ une base de $H^1_f(M)$, $\omega_{f,M^*(1)}$ une base de $H^1_f(M^*(1))$, ω_M^0 une base de $H^0(M)$ et $\omega^0_{M^*(1)}$ une base de $H^0(M^*(1))$. Le nombre

$$\frac{[\omega_B^+][\operatorname{coker}\alpha_M]}{[\omega_{\mathrm{tg}}][\ker \alpha_M]} \cdot \frac{[\ker \alpha_M][\omega_{f,M}][\omega_{f,M^*(1)}]}{[\omega_M^0][\omega^0_{M^*(1)}][\operatorname{coker}\alpha_M]}$$

ne dépend ni du choix des bases de $\ker \alpha_M$ et $\operatorname{coker}\alpha_M$ ni, à multiplication près par un élément de $\mathbf{Q}^*$, du choix des bases ω_{tg}, ω_B^+, $\omega_{f,M}$, $\omega_{f,M^*(1)}$, ω_M^0 et $\omega^0_{M^*(1)}$; il sera noté $\Omega_\infty(M)$.

CONJECTURE 2.1 (Deligne-Beilinson). — *$L(M,s)$ a, en $s=0$, un zéro d'ordre $r_\infty(M) = \dim_{\mathbf{Q}} H^1_f(M^*(1)) - \dim_{\mathbf{Q}} H^0(M^*(1))$ et $L^*(M,0) = \lim_{s\to 0} s^{-r_\infty(M)} L(M,s)$ est égal, à multiplication près par un élément de $\mathbf{Q}^*$, à $\Omega_\infty(M)$.*

2.3. **Les anneaux de Fontaine**

Tate a montré [52] qu'il n'existait pas dans $\mathbf{C}_p$ d'analogue p-adique de $2i\pi$ et donc par conséquent que les périodes p-adiques des variétés algébriques ne pouvaient pas vivre dans $\mathbf{C}_p$. Ceci a conduit Fontaine [16, 17] à construire des anneaux $\mathbf{B}_{\mathrm{cris}} \subset \mathbf{B}_{\mathrm{dR}}$ un peu compliqués mais fort utiles.

$\mathbf{B}_{\mathrm{dR}}$ est un corps contenant les périodes de toutes les variétés algébriques. C'est le corps des fractions d'un anneau $\mathbf{B}^+_{\mathrm{dR}}$ obtenu en complétant $\overline{\mathbf{Q}}_p$ pour une topologie plus fine que la topologie p-adique faisant intervenir "les dérivées successives des éléments de $\overline{\mathbf{Q}}_p$ vus comme fonctions algébriques de p". Ceci fait de $\mathbf{B}^+_{\mathrm{dR}}$ un anneau topologique muni d'une action continue de $\mathscr{G}_{\mathbf{Q}_p}$. Il contient en particulier $\overline{\mathbf{Q}}_p$ et un analogue p-adique de $2i\pi$ que nous noterons t dans la suite et sur lequel $g \in \mathscr{G}_{\mathbf{Q}_p}$ agit par $g(t) = \chi_{\mathrm{cycl}}(g)t$. D'autre part, $\mathbf{B}^+_{\mathrm{dR}}$ est un anneau de valuation discrète complet dont l'idéal maximal est engendré par t et le corps résiduel est $\mathbf{C}_p$ (ce qui explique que l'on ne voit pas t dans $\mathbf{C}_p$). On munit $\mathbf{B}_{\mathrm{dR}} = \mathbf{B}^+_{\mathrm{dR}}[\frac{1}{t}]$ de la filtration définie par $\mathbf{B}^i_{\mathrm{dR}} = t^i \mathbf{B}^+_{\mathrm{dR}}$ si $i \in \mathbf{Z}$. Finalement, on a $\mathbf{B}_{\mathrm{dR}}^{\mathscr{G}_{\mathbf{Q}_p}} = \mathbf{Q}_p$.

$\mathbf{B}_{\mathrm{cris}}$ est un sous-anneau de $\mathbf{B}_{\mathrm{dR}}$ stable par $\mathscr{G}_{\mathbf{Q}_p}$ qui contient les périodes des variétés algébriques ayant bonne réduction modulo p. C'est un anneau topologique muni d'une action d'un morphisme de Frobenius φ commutant à l'action de $\mathscr{G}_{\mathbf{Q}_p}$. L'intersection de $\mathbf{B}_{\mathrm{cris}}$ avec $\overline{\mathbf{Q}}_p$ est l'extension maximale non ramifiée $\mathbf{Q}_p^{\mathrm{nr}}$ de $\mathbf{Q}_p$ et t est un élément de $\mathbf{B}_{\mathrm{cris}}$ sur lequel φ agit par multiplication par p. D'autre part, les anneaux $\mathbf{B}_{\mathrm{cris}}$ et $\mathbf{B}_{\mathrm{dR}}$ sont

reliés par les suites exactes fondamentales

$$0 \to \mathbf{Q}_p \to \mathbf{B}_{\text{cris}}^{\varphi=1} \to \mathbf{B}_{\text{dR}}/\mathbf{B}_{\text{dR}}^{+} \to 0, \tag{2}$$

$$0 \to \mathbf{Q}_p \to \mathbf{B}_{\text{cris}} \to \mathbf{B}_{\text{cris}} \oplus \mathbf{B}_{\text{dR}}/\mathbf{B}_{\text{dR}}^{+} \to 0, \tag{3}$$

où, dans la seconde suite exacte, l'application de $\mathbf{B}_{\text{cris}}$ dans $\mathbf{B}_{\text{cris}} \oplus \mathbf{B}_{\text{dR}}/\mathbf{B}_{\text{dR}}^{+}$ est celle qui à x associe $((1-\varphi)x, x)$.

2.4. **Représentations p-adiques attachées aux motifs**

2.4.1. *La hiérarchie des représentations p-adiques.* Si V est une représentation p-adique de $\mathscr{G}_{\mathbf{Q}_p}$, on pose $\mathrm{D}_{\text{dR}}(V) = (\mathbf{B}_{\text{dR}} \otimes V)^{\mathscr{G}_{\mathbf{Q}_p}}$ et $\mathrm{D}_{\text{cris}}(V) = (\mathbf{B}_{\text{cris}} \otimes V)^{\mathscr{G}_{\mathbf{Q}_p}}$. Ceci fait de $\mathrm{D}_{\text{dR}}(V)$ un $\mathbf{Q}_p$-espace vectoriel que l'on munit de la filtration décroissante définie par $\mathrm{D}_{\text{dR}}^{i}(V) = (\mathbf{B}_{\text{dR}}^{i} \otimes V)^{\mathscr{G}_{\mathbf{Q}_p}}$ et $\mathrm{D}_{\text{cris}}(V) \subset \mathrm{D}_{\text{dR}}(V)$ est de plus muni d'une action de φ. On a $\dim_{\mathbf{Q}_p} \mathrm{D}_{\text{cris}}(V) \leq \dim_{\mathbf{Q}_p} \mathrm{D}_{\text{dR}}(V) \leq \dim_{\mathbf{Q}_p} V$ et on dit que V est cristalline si $\dim_{\mathbf{Q}_p} \mathrm{D}_{\text{cris}}(V) = \dim_{\mathbf{Q}_p} V$ et de de Rham si $\dim_{\mathbf{Q}_p} \mathrm{D}_{\text{dR}}(V) = \dim_{\mathbf{Q}_p} V$.

(Mot 6) Si M est un motif, la représentation $M_{B,p}$ est de de Rham et on dispose d'un isomorphisme "périodes p-adiques" $\iota_{M,p} : \mathrm{D}_{\text{dR}}(M_{B,p}) \to M_{\text{dR},p} = \mathbf{Q}_p \otimes M_{\text{dR}}$ de $\mathbf{Q}_p$-espaces vectoriels filtrés. De plus, le déterminant de $1 - X\varphi$ agissant sur $\mathrm{D}_{\text{cris}}(M_{B,p})$ est égal à $E_p(M, X)$; en particulier, si M a bonne réduction en p, alors $M_{B,p}$ est cristalline.

Réciproquement, on conjecture [20, 27] que $\mathbf{B}_{\text{dR}}$ est juste assez fin pour faire le tri entre ce qui provient de la géométrie et ce qui n'en provient pas, ce qui se traduit par les conjectures 2.2 et 2.3 ci-dessous.

CONJECTURE 2.2 (Fontaine-Mazur). — *Si V est une représentation p-adique irréductible de $\mathscr{G}_{\mathbf{Q}}$ qui est non ramifiée en dehors d'un ensemble fini de places et dont la restriction à un groupe de décomposition en p est de de Rham, alors V est la représentation p-adique associée à un motif pur.*

2.4.2. *Extensions de motifs et de représentations p-adiques.* Bloch et Kato ont introduit [9] des sous-groupes $H_e^1(\mathbf{Q}_p, V) \subset H_f^1(\mathbf{Q}_p, V) \subset H_g^1(\mathbf{Q}_p, V)$ de $H^1(\mathbf{Q}_p, V)$. Ce sont les noyaux respectifs des applications naturelles de $H^1(\mathbf{Q}_p, V)$ dans $H^1(\mathbf{Q}_p, \mathbf{B}_{\text{cris}}^{\varphi=1} \otimes V)$, $H^1(\mathbf{Q}_p, \mathbf{B}_{\text{cris}} \otimes V)$ et $H^1(\mathbf{Q}_p, \mathbf{B}_{\text{dR}} \otimes V)$. La signification de ces sous-groupes est la suivante. Si $x \in H^1(\mathbf{Q}_p, V)$, l'extension V_x de $\mathbf{Q}_p$ par V qui lui est associée est cristalline (resp. de de Rham) si et seulement si V est cristalline (resp. de de Rham) et $x \in H_f^1(\mathbf{Q}_p, V)$ (resp. $x \in H_g^1(\mathbf{Q}_p, V)$). Rappelons d'autre part que le cup-produit fournit un accouplement parfait

$$H^1(\mathbf{Q}_p, V) \times H^1(\mathbf{Q}_p, V^*(1)) \longrightarrow H^2(\mathbf{Q}_p, \mathbf{Q}_p(1)) \cong \mathbf{Q}_p$$

qui permet en particulier de voir $H^1(\mathbf{Q}_p, V^*(1))$ comme le dual de $H^1(\mathbf{Q}_p, V)$ et que, dans cette dualité, les orthogonaux respectifs de $H_e^1(\mathbf{Q}_p, V)$, $H_f^1(\mathbf{Q}_p, V)$ et $H_g^1(\mathbf{Q}_p, V)$ sont $H_g^1(\mathbf{Q}_p, V^*(1))$, $H_f^1(\mathbf{Q}_p, V^*(1))$ et $H_e^1(\mathbf{Q}_p, V^*(1))$.

(Mot 7) Si $x \in H^1(M)$ a bonne réduction en p, alors $x_{B,p} \in H^1_f(\mathbf{Q}_p, M_{B,p})$.

Si $\ell \neq p$ et V est une représentation p-adique de $\mathscr{G}_{\mathbf{Q}_\ell}$, notons $H^1_f(\mathbf{Q}_\ell, V)$ le sous-groupe des éléments de $H^1(\mathbf{Q}_\ell, V)$ dont la restriction au sous-groupe d'inertie I_ℓ est triviale.

Soit S un ensemble fini de nombres premiers et soit $\mathscr{G}_{\mathbf{Q},S}$ le groupe de Galois de l'extension maximale de $\mathbf{Q}$ non ramifiée en dehors de l'infini et de S. Si V est une représentation p-adique de $\mathscr{G}_{\mathbf{Q},S}$ et $i \in \mathbf{N}$, notons $H^i(V)$ le groupe $H^i(\mathscr{G}_{\mathbf{Q},S}, V)$. Notons $H^1_g(V)$ le sous-groupe des éléments x de $H^1(V)$ dont la restriction à $\mathscr{G}_{\mathbf{Q}_p}$ appartient à $H^1_g(\mathbf{Q}_p, V)$. Si Σ est un sous-ensemble de S, notons $H^1_{f,\Sigma}(V)$ le sous-groupe des éléments x de $H^1_g(V)$ dont la restriction à $\mathscr{G}_{\mathbf{Q}_\ell}$ appartient à $H^1_f(\mathbf{Q}_\ell, V)$ si $\ell \notin \Sigma$ et $H^1_f(V)$ le groupe correspondant à $\Sigma = \emptyset$.

La conjecture suivante [27] est l'analogue pour les extensions de motifs de la conjecture de Fontaine-Mazur.

CONJECTURE 2.3 (Jannsen). — *Soit S un ensemble fini de nombres premiers contenant p. Si M est un motif ayant bonne réduction en dehors de S, les applications naturelles de $\mathbf{Q}_p \otimes H^1_f(M)$ et $\mathbf{Q}_p \otimes H^1_{f,S}(M)$ dans $H^1_f(M_{B,p})$ et $H^1_{f,S}(M_{B,p})$ respectivement sont des isomorphismes.*

Si M est le motif associé au H^1 d'une variété abélienne X, la conjecture ci-dessus est plus ou moins équivalente à la finitude de la p-partie du groupe de Chafarevitch-Tate de X. Finalement, on a la conjecture de Tate

CONJECTURE 2.4 (Tate). — *Si M est un motif ayant bonne réduction en dehors de S, l'application naturelle de $\mathbf{Q}_p \otimes H^0(M)$ dans $H^0(M_{B,p})$ est un isomorphisme.*

2.5. La conjecture de Bloch-Kato

2.5.1. *Le groupe de Chafarevitch-Tate.* Soient V une représentation p-adique de $\mathscr{G}_{\mathbf{Q},S}$ et T un réseau de V stable par $\mathscr{G}_{\mathbf{Q},S}$. On définit les groupes $H^1_f(T)$, $H^1_{f,S}(T)$ et $H^1_f(\mathbf{Q}_\ell, T)$ comme les images inverses respectives de $H^1_f(V)$, $H^1_{f,S}(V)$ et $H^1_f(\mathbf{Q}_\ell, V)$ dans $H^1(T)$ et $H^1(\mathbf{Q}_\ell, T)$. Si ℓ est une place finie, on définit $H^1_f(\mathbf{Q}_\ell, V/T)$ comme l'image de $H^1_f(\mathbf{Q}_\ell, V)$ dans $H^1(\mathbf{Q}_\ell, V/T)$ et $H^1_f(V/T)$ comme l'ensemble des éléments de $H^1(V/T)$ dont l'image dans $H^1(\mathbf{Q}_\ell, V/T)$ appartient à $H^1_f(\mathbf{Q}_\ell, V/T)$ pour tout nombre premier ℓ et est nulle si $\ell = \infty$. Finalement, on définit le groupe $Ш(T)$ comme le quotient de $H^1_f(V/T)$ par $H^1_f(V)$; c'est un groupe fini (dans le cas d'une variété abélienne X, c'est le quotient de la p-partie du groupe de Chafarevitch-Tate usuel de X par sa partie divisible).

2.5.2. *Nombres de Tamagawa.* Soient V une représentation p-adique de $\mathscr{G}_{\mathbf{Q}_\ell}$ et T un réseau de V stable par $\mathscr{G}_{\mathbf{Q}_\ell}$.

Si $\ell = \infty$, on pose $\mathrm{Tam}^0_\ell(T) = |H^1(\mathbf{R}, T)|$ et si ℓ est un nombre premier distinct de p, on définit $\mathrm{Tam}^0_\ell(T)$ comme le cardinal du sous-groupe de torsion de $H^1(I_\ell, T)^{\mathscr{G}_{\mathbf{Q}_\ell}}$. En particulier, on a $\mathrm{Tam}^0_\ell(T) = 1$ si V n'est pas ramifiée en ℓ.

Si $\ell = p$, on dispose, en considérant la suite exacte de cohomologie obtenue en tensorisant la suite exacte fondamentale (3) par V, de la suite exacte

$$0 \to H^0(\mathbf{Q}_p, V) \to \mathrm{D}_{\mathrm{cris}}(V) \to \mathrm{D}_{\mathrm{cris}}(V) \oplus t_V \to H^1_f(\mathbf{Q}_p, V) \to 0,$$

ce qui nous permet, si ω_{tg} est une base de t_V et D est un réseau de $\mathrm{D}_{\mathrm{cris}}(V)$, de poser $\mathrm{Tam}^0_{p,\omega_{\mathrm{tg}}}(T) = \frac{[H^0(\mathbf{Q}_p,T)][D\oplus\omega_{\mathrm{tg}}]}{[D][H^1_f(\mathbf{Q}_p,T)]}$ et cette quantité ne dépend pas du choix de D (mais elle fait intervenir le déterminant de $1-\varphi$ sur $D/D^{\varphi=1}$).

Finalement, on pose

$$\mathrm{Tam}^0_{\omega_{\mathrm{tg}}}(T) = \mathrm{Tam}^0_{p,\omega_{\mathrm{tg}}}(T) \prod_{\ell \neq p} \mathrm{Tam}^0_\ell(T).$$

2.5.3. *Énoncé de la conjecture.* On reprend les notations de la conjecture de Deligne-Beilinson et on suppose que M a bonne réduction en dehors de S. Modulo la conjecture de Tate et celle de Jannsen, les $\mathbf{Z}_p$-modules engendrés par $\omega_{f,V}$, $\omega_{f,V^*(1)}$, ω^0_V et $\omega^0_{V^*(1)}$ sont des réseaux de $H^1_f(M_{B,p})$, $H^1_f(M^*(1)_{B,p})$, $H^0(M_{B,p})$ et $H^0(M^*(1)_{B,p})$ respectivement. Soit T un réseau de M_B tel que $T_p = \mathbf{Z}_p \otimes T$ soit un sous-réseau de $M_{B,p}$ stable par $\mathscr{G}_{\mathbf{Q},S}$ et prenons pour ω_B^+ une base de T^+.

CONJECTURE 2.5 (Bloch-Kato). — *Si $x \sim_p y$ signifie que x et y ont même valuation p-adique, alors*

$$\frac{L^*(M,0)}{\Omega_\infty(M)} \cdot \frac{[H^1_f(T_p) : \omega_{f,V}][H^1_f(T_p^*(1)) : \omega_{f,V^*(1)}]}{[H^0(T_p) : \omega^0_V][H^0(T_p^*(1)) : \omega^0_{V^*(1)}]} \sim_p \mathrm{Tam}^0_{\omega_{\mathrm{tg}}}(T_p)|Ш(T_p^*(1))|.$$

2.6. L'exponentielle de Bloch-Kato et sa duale

Si V est une représentation de de Rham de $\mathscr{G}_{\mathbf{Q}_p}$, tensorisant la suite exacte fondamentale (2) avec V et prenant la suite exacte de cohomologie associée, on obtient la suite exacte

$$0 \to V^{\mathscr{G}_{\mathbf{Q}_p}} \to \mathrm{D}^{\varphi=1}_{\mathrm{cris}}(V) \to t_V \to H^1_e(\mathbf{Q}_p, V) \to 0,$$

où l'on a posé $t_V = \mathrm{D}_{\mathrm{dR}}(V)/\mathrm{D}^0_{\mathrm{dR}}(V)$. On appelle exponentielle de Bloch-Kato l'application $\exp_V$ de connexion $t_V \to H^1_e(\mathbf{Q}_p, V)$ que l'on verra aussi souvent comme une application de $\mathrm{D}_{\mathrm{dR}}(V)$ dans $H^1(\mathbf{Q}_p, V)$. Si $k \gg 0$, alors $\mathrm{D}^0_{\mathrm{dR}}(V(k)) = 0$ et $\exp_{V(k)}$ est un isomorphisme de $\mathrm{D}_{\mathrm{dR}}(V(k))$ sur $H^1(\mathbf{Q}_p, V(k))$.

Si V est une représentation de $\mathscr{G}_{\mathbf{Q}_p}$, l'accouplement $\langle\ ,\ \rangle_{\mathrm{dR}}$ obtenu en composant les applications

$$\mathrm{D}_{\mathrm{dR}}(V) \times D_{\mathrm{dR}}(V^*(1)) \to \mathrm{D}_{\mathrm{dR}}(V \otimes V^*(1)) \to \mathrm{D}_{\mathrm{dR}}(\mathbf{Q}_p(1)) = t^{-1}\mathbf{Q}_p \cong \mathbf{Q}_p$$

est non dégénéré, ce qui fait que $\mathrm{D}_{\mathrm{dR}}(V^*(1))$ s'identifie naturellement au dual de $\mathrm{D}_{\mathrm{dR}}(V)$ et $\mathrm{D}^0_{\mathrm{dR}}(V)$ s'identifie au dual de $t_{V^*(1)}$. Ceci permet de voir l'application $\exp^*_{V^*(1)}$ transposée de l'application $\exp_{V^*(1)} : t_{V^*(1)} \to H^1(\mathbf{Q}_p, V^*(1))$ comme une application de $H^1(\mathbf{Q}_p, V)$ dans $\mathrm{D}^0_{\mathrm{dR}}(V)$. Si $k \gg 0$, l'application $\exp^*_{V^*(1+k)}$ est un isomorphisme de $H^1(\mathbf{Q}_p, V(-k))$

sur $D_{\mathrm{dR}}(V(-k))$. Cette définition de l'application exponentielle duale n'est pas très commode pour les calculs (il faut dualiser deux fois), mais on a la proposition suivante [28].

PROPOSITION 2.6 (Kato). — *Si V est une représentation de de Rham, l'application qui à $x \in \mathrm{D}_{\mathrm{dR}}(V)$ associe le cocycle $\tau \to x \log_p \chi_{\mathrm{cycl}}(\tau) \in \mathrm{D}_{\mathrm{dR}}(V) \subset \mathbf{B}_{\mathrm{dR}} \otimes V$ induit un isomorphisme de $\mathrm{D}_{\mathrm{dR}}(V)$ sur $H^1(\mathbf{Q}_p, \mathbf{B}_{\mathrm{dR}} \otimes V)$ et l'application $\exp^*_{V^*(1)}$ est la composée de l'inverse de cet isomorphisme avec l'application naturelle $H^1(\mathbf{Q}_p, V) \to H^1(\mathbf{Q}_p, \mathbf{B}_{\mathrm{dR}} \otimes V)$.*

2.7. **Hauteurs p-adiques**

La construction de l'accouplement "hauteur p-adique" décrit ci-dessous est tirée de [36, 39] et [41, Chap. III]; c'est une adaptation de la construction de l'accouplement complexe via la théorie des biextensions [46, 47].

Si $\ell \neq p$ et $x \in \mathbf{Q}_\ell^*$, alors $|x|_\ell \in \mathbf{Q}^*$ et $\log_p |x|_\ell$ est un élément bien défini de $\mathbf{Q}_p$. D'autre part, d'après la théorie de Kummer, si K est un corps de caractéristique 0, alors $H^1(K, \mathbf{Z}_p(1))$ est le complété p-adique de K^*, ce qui permet de prolonger par continuité l'application $x \to -\log_p |x|_\ell$ en une application $\mathbf{Q}_p$-linéaire $\mathrm{inv}_\ell : H^1(\mathbf{Q}_\ell, \mathbf{Q}_p(1)) \to \mathbf{Q}_p$. On prolonge de même l'application $\log_p : \mathbf{Q}_p^* \to \mathbf{Q}_p$ en une application $\mathbf{Q}_p$-linéaire $\mathrm{inv}_p : H^1(\mathbf{Q}_p, \mathbf{Q}_p(1)) \to \mathbf{Q}_p$. Comme $\log_p p = 0$ par définition, la formule du produit pour les nombres rationnels montre que si $y \in H^1(\mathbf{Q}_p(1))$, alors $\sum \mathrm{inv}_v(y) = 0$.

Soit V une représentation p-adique de $\mathscr{G}_{\mathbf{Q},S}$ dont la restriction à $\mathscr{G}_{\mathbf{Q}_p}$ est de de Rham. On suppose de plus que $\mathrm{D}_{\mathrm{cris}}(V)^{\varphi=1} = \mathrm{D}_{\mathrm{cris}}(V)^{\varphi=p^{-1}} = 0$, ce qui implique en particulier que les applications $\exp_V : t_V \to H^1_f(\mathbf{Q}_p, V)$ et $\exp_{V^*(1)} : t_{V^*(1)} \to H^1_f(\mathbf{Q}_p, V^*(1))$ sont des isomorphismes. Leurs inverses respectifs seront notés $\log_V$ et $\log_{V^*(1)}$.

Soit $x \in H^1_f(V^*(1))$. Il lui correspond une extension V_x de V par $\mathbf{Q}_p(1)$ et comme nous avons supposé $\mathrm{D}_{\mathrm{cris}}(V)^{\varphi=p^{-1}} = 0$, la suite exacte $0 \to \mathrm{D}_{\mathrm{dR}}(V^*(1)) \to \mathrm{D}_{\mathrm{dR}}(V_x^*(1)) \to \mathbf{Q}_p \to 0$ admet un unique scindage σ_x^* tel que $\sigma_x^*(\mathbf{Q}_p) = \mathrm{D}_{\mathrm{cris}}(V_x^*(1))^{\varphi=1}$. Par dualité, cela nous fournit une application $\sigma_x : \mathrm{D}_{\mathrm{dR}}(V) \to \mathrm{D}_{\mathrm{dR}}(V_x)$ scindant la suite exacte $0 \to \mathrm{D}_{\mathrm{dR}}(\mathbf{Q}_p(1)) \to \mathrm{D}_{\mathrm{dR}}(V_x) \to \mathrm{D}_{\mathrm{dR}}(V) \to 0$. Cette application permet, modulo le choix d'un scindage $s_{\mathrm{Hdg}} : t_V \to \mathrm{D}_{\mathrm{dR}}(V)$ de la filtration de Hodge de définir une section $s_{\mathbf{Q}_p,x}$ de l'application (surjective) de $H^1_f(\mathbf{Q}_p, V_x)$ sur $H^1_f(\mathbf{Q}_p, V)$ en composant les flèches passant par le bas du diagramme ci-dessous

$$\begin{array}{ccccc} H^1_f(\mathbf{Q}_p, V) & & \xrightarrow{\quad s_{\mathbf{Q}_p,x} \quad} & & H^1_f(\mathbf{Q}_p, V_x) \\ \downarrow{\scriptstyle \log_V} & & & & \uparrow{\scriptstyle \exp_{V_x}} \\ t_V & \xrightarrow{s_{\mathrm{Hdg}}} & \mathrm{D}_{\mathrm{dR}}(V) & \xrightarrow{\sigma_x} & \mathrm{D}_{\mathrm{dR}}(V_x) \end{array}$$

Soit maintenant $y \in H^1_f(V)$. Choisissons $s_{\mathbf{Q},x}(y) \in H^1(V_x)$ relevant y. Par construction, l'image de $s_{\mathbf{Q},x}(y) - s_{\mathbf{Q}_p,x}(y)$ dans $H^1(\mathbf{Q}_p, V)$ est nulle et donc $s_{\mathbf{Q},x}(y) - s_{\mathbf{Q}_p,x}(y) \in H^1(\mathbf{Q}_p, \mathbf{Q}_p(1))$. Si $\ell \neq p$, choisissons un relèvement $s_{\mathbf{Q}_\ell,x}(y)$ dans $H^1_f(\mathbf{Q}_\ell, V_x)$, ce qui fait que l'on a aussi $s_{\mathbf{Q},x}(y) - s_{\mathbf{Q}_\ell,x}(y) \in H^1(\mathbf{Q}_\ell, \mathbf{Q}_p(1))$. Comme on ne peut changer $s_{\mathbf{Q},x}(y)$

que par un élément de $H^1(\mathbf{Q}_p(1))$, la somme $\sum_\ell \mathrm{inv}_\ell(s_{\mathbf{Q},x}(y) - s_{\mathbf{Q}_\ell,x}(y))$ ne dépend que de x et y (et du choix de s_{Hdg}) et l'application "hauteur p-adique" $\langle\ ,\ \rangle_{\mathrm{ht}} : H^1_f(V^*(1)) \times H^1_f(V)$ ainsi définie est bilinéaire. On étend $\langle\ ,\ \rangle_{\mathrm{ht}}$ en un accouplement

$$\langle\ ,\ \rangle_{\mathrm{ht}} : \left(H^1_f(V^*(1)) \oplus t_V^*\right) \times \left(H^1_f(V) \oplus \mathrm{D}_{\mathrm{dR}}(V)\right) \to \mathbf{Q}_p$$

grâce à la formule

$$\langle (x,\omega),(y,v)\rangle_{\mathrm{ht}} = \langle x,y\rangle_{\mathrm{ht}} + \langle \omega, \overline{v} - \log_V(y)\rangle_{\mathrm{dR}} + \langle \log_{V^*(1)}(x), v - s_{\mathrm{Hdg}}(\overline{v})\rangle_{\mathrm{dR}},$$

où $\overline{v}$ désigne l'image de $v \in \mathrm{D}_{\mathrm{dR}}(V)$ dans t_V et $\langle\ ,\ \rangle_{\mathrm{dR}}$ provient de la dualité entre $\mathrm{D}_{\mathrm{dR}}(V^*(1))$ et $\mathrm{D}_{\mathrm{dR}}(V)$.

Si $s : t_V \to \mathrm{D}_{\mathrm{dR}}(V)$ est un autre scindage de la filtration de Hodge, il existe une application $\mathbf{Q}_p$-linéaire $u : t_V \to \mathrm{D}^0_{\mathrm{dR}}(V)$ telle que l'on ait $s = s_{\mathrm{Hdg}} + u$. Par dualité, on obtient une application u^* de $\mathrm{D}^0_{\mathrm{dR}}(V)^* = t_{V^*(1)}$ dans t_V et on montre que l'accouplement $\langle\ ,\ \rangle_u$ obtenu à partir de s est relié à $\langle\ ,\ \rangle_{\mathrm{ht}}$ par la formule

$$\langle (x,\omega),(y,v)\rangle_u = \langle (x, \omega + u^*(\log_{V^*(1)}(x))),(y,v)\rangle_{\mathrm{ht}}. \tag{4}$$

2.8. **Conjecture de Beilinson p-adique**

Soit M un motif de dimension d défini sur $\mathbf{Q}$ et ayant bonne réduction en p. La fonction L p-adique de M va être "définie" par interpolation p-adique des valeurs $\frac{L^*(M,n)}{\Omega_\infty(M(n))}$ pour $n \in \mathbf{Z}$ (ne pas oublier que l'on a $L^*(M,n) = L^*(M(n),0)$). On pose $d^+ = d^+(M)$ et $d^- = d^-(M)$.

Si $x \in M_{\mathrm{dR},p}$, on notera $t^{-n}x$ son image dans $M(n)_{\mathrm{dR},p}$ par l'isomorphisme canonique. Cette notation est compatible avec les isomorphismes

$$M(n)_{\mathrm{dR},p} \cong \mathrm{D}_{\mathrm{dR}}(M(n)_{B,p}) = t^{-n}\mathrm{D}_{\mathrm{dR}}(M_{B,p}) \cong t^{-n}M_{\mathrm{dR},p}$$

et reflète le fait que l'action de φ sur $M(n)_{\mathrm{dR},p} = \mathrm{D}_{\mathrm{cris}}(M(n)_{B,p})$ se déduit de celle sur $M_{\mathrm{dR},p}$ en la multipliant par p^{-n}.

On n'a pas besoin de toutes les valeurs de n pour assurer l'unicité de la fonction L p-adique, aussi nous supposerons que ni 1 ni p^{-1} ne sont valeurs propres de φ sur $M(n)_{\mathrm{dR},p}$. En particulier cela implique que $H^0(\mathbf{Q}_p, M(n)_{B,p}) = H^0(\mathbf{Q}_p, M^*(1-n)_{B,p}) = 0$ et donc que $H^0(M(n)) = H^0(M^*(1-n)) = 0$.

Reprenons les bases $\omega_{\mathrm{tg},n}$ de $t_{M(n)}$, $\omega_{f,M(n)}$ de $H^1_f(M(n))$, et $\omega_{f,M^*(1-n)}$ de $H^1_f(M^*(1-n))$ qui nous ont servi à définir $\Omega_\infty(M(n))$. Si $v = (v_1,\dots,v_d)$ est une famille d'éléments de $M_{\mathrm{dR},p}$, soient $v^+ = (v_1,\dots,v_{d^+})$ et $v^- = (v_{d^+1},\dots,v_d)$. On définit alors la période p-adique de $M(n)$ par la formule

$$\Omega_p(M(n),v) = \left\langle \omega_{f,M^*(1-n)} \oplus \omega_{\mathrm{tg}}, \omega_{f,M(n)} \oplus v^{(-1)^n} \right\rangle_{\mathrm{ht}},$$

la formule (4) montrant que ce déterminant est indépendant du choix du scindage de la filtration de Hodge ayant permis de définir $\langle\ ,\ \rangle_{\mathrm{ht}}$. Remarquons d'autre part que $\Omega_p(M(n),v)$ ne dépend, si n est pair (resp. impair), que de l'image de v dans $\wedge^{d^+(M(n))}\mathrm{D}_{\mathrm{dR}}(V)$ par

l'application qui, à la famille $(v_1, \dots, v_d)$, associe le produit extérieur de ses $d^+(M(n))$ premiers (resp. derniers) vecteurs et qu'elle n'en dépend que de manière linéaire.

Soit $\omega_B = (\omega_B^+, \omega_B^-)$, où ω_B^+ est une base de M_B^+ et ω_B^- est une base de M_B^-. Si $n \in \mathbf{N}$, prenons $\omega_B^{(-1)^n}$ pour base de $M(n)_B^+$ pour faire le calcul de $\Omega_\infty(M(n))$. Finalement, si $n \leq 0$ (resp. $n \geq 1$), posons $\Gamma^*(n) = \lim_{s\to 0} s\Gamma(n+s) = \frac{(-1)^n}{n!}$ (resp. $\Gamma^*(n) = \Gamma(n) = (n-1)!$).

CONJECTURE 2.7 (Perrin-Riou). — *Si v est une famille de d éléments de $M_{\mathrm{dR},p}$, il existe une (unique) pseudo-distribution tempérée $\mu_{M,v}$ sur Γ telle que, si $L_p(M \otimes \chi_{\mathrm{cycl}}^n, v, s) = \int_\Gamma \chi_{\mathrm{cycl}}^n (\chi_{\mathrm{cycl}})^s \mu_{M,v}$ et si ni 1 ni p^{-1} ne sont valeurs propres de φ sur $M(n)_{\mathrm{dR},p}$, alors la fonction $L_p(M \otimes \chi_{\mathrm{cycl}}^n, v, s)$ a, en $s = 0$, un zéro d'ordre au moins égal à*

$$r_p(M(n)) = \begin{cases} \dim_{\mathbf{Q}} H^1_f(M^*(1-n)) & \text{si } n \geq 1 \\ \dim_{\mathbf{Q}} H^1_f(M^*(1-n)) - d^+(M(n)) & \text{si } n \leq 0 \end{cases}$$

et si on pose $L_p^*(M \otimes \chi_{\mathrm{cycl}}^n, v) = \lim_{s\to 0} s^{-r_p(M(n))} L_p(M \otimes \chi_{\mathrm{cycl}}^n, v, s)$, alors

$$L_p^*(M \otimes \chi_{\mathrm{cycl}}^n, v) = \frac{L^*(M,n) E_p(M, p^{-n})}{\Omega_\infty(M(n))} \Omega_p\Big(M(n), \frac{1 - p^{-1}\varphi^{-1}}{1-\varphi}\Big(\frac{\Gamma^*(n)}{(-t)^n} v\Big)\Big).$$

REMARQUE 2.8. — (i) Sous cette forme, la conjecture semble totalement hors d'atteinte mais si on sait par ailleurs construire la fonction L p-adique de M, elle fournit un analogue p-adique de la conjecture de Beilinson.

(ii) La fonction L p-adique attachée à M (i.e. la fonction $\psi \to L_p(M \otimes \psi, v) = \int_\Gamma \psi\, \mu_{M,v}$) n'est bien définie, si elle existe, qu'à multiplication près par un élément de $\mathbf{Q}^* \times \mathbf{Q}^*$ (un facteur pour les caractères pairs et un autre pour les caractères impairs). D'autre part, elle dépend de manière linéaire de $([v^+], [v^-])$ et $v = (v^+, v^-)$ doit être considéré comme un paramètre supplémentaire jouant le rôle de (M_B^+, M_B^-).

(iii) On peut donner (cf. [41]) une formule pour $L_p(M \otimes \eta\chi_{\mathrm{cycl}}^n, v)$ pour tout caractère d'ordre fini η et pour tout entier n tel que ni 1 ni p^{-1} ne soit valeur propre de φ sur $M(n)_{\mathrm{dR},p}$.

(iv) L'opérateur $\frac{1-p^{-1}\varphi^{-1}}{1-\varphi}$ doit être interprété comme un facteur d'Euler en p que l'on doit rajouter pour rendre le terme de droite continu p-adiquement. En particulier, si v est vecteur propre de φ, la formule d'interpolation se simplifie et on voit apparaître un facteur d'Euler au sens usuel du terme. Plus exactement, comme on a déjà enlevé le facteur d'Euler au-dessus de p en multipliant par $E_p(M, p^{-n})$, le terme $(1-\varphi)^{-1}$ réintroduit une partie du facteur d'Euler de $M(n)$ alors que le terme $(1 - p^{-1}\varphi^{-1})$ enlève une partie du facteur d'Euler de $M^*(1-n)$. Le cas de la fonction zêta de Riemann est assez éclairant pour comprendre l'effet de $\frac{1-p^{-1}\varphi^{-1}}{1-\varphi}$. Elle correspond au motif trivial $\mathbf{Q}(0)$ et on a $t_{\mathbf{Q}(n)} = \mathbf{Q}$ si $n \geq 1$ et $t_{\mathbf{Q}(n)} = 0$ si $n \leq 0$. On prend 1 comme base de $\mathbf{Q}(0)_{\mathrm{dR}} \subset \mathbf{Q}(0)_{\mathrm{dR},p}$ et $(1,0)$ pour ω_B. D'autre part, on a $d^+(\mathbf{Q}(n)) = 1$ si n est pair et $d^+(\mathbf{Q}(n)) = 0$ si n est impair. Finalement, si $n \geq 1$ est pair ou si $n \leq 0$ est impair, on a $H^0(\mathbf{Q}(n)) = H^0(\mathbf{Q}(1-n)) = H^1(\mathbf{Q}(n)) = H^1(\mathbf{Q}(1-n)) = 0$, ce qui fait que si $n \leq 0$ est impair,

tous les espaces vectoriels intervenant dans les définitions de $\Omega_\infty(\mathbf{Q}(n))$ et $\Omega_p(\mathbf{Q}(n),1)$ sont nuls et ces périodes valent donc 1 toutes les deux et la valeur que l'on doit interpoler est $\zeta(n)E_p(\mathbf{Q}(0),p^{-n}) = \zeta(n)(1-p^{-n})$. Dans le cas $n \geq 1$ pair, la période $\Omega_\infty(\mathbf{Q}(n))$ correspond à l'isomorphisme $\mathbf{Q}(n)_B^+ \otimes \mathbf{R} \to t_{\mathbf{Q}(n)} \otimes \mathbf{R}$ et vaut $(2i\pi)^n$. D'un autre côté, on a $\frac{1-p^{-1}\varphi^{-1}}{1-\varphi}\frac{\Gamma^*(n)}{(-t)^n} = \Gamma(n)\frac{1-p^{n-1}}{1-p^{-n}}(-t)^{-n}$, ce qui fait que la valeur que l'on interpole doit être, d'après la conjecture

$$\frac{\zeta(n)E_p(\mathbf{Q}(0),p^{-n})}{(2i\pi)^n}\cdot\frac{\Gamma(n)(1-p^{n-1})}{1-p^{-n}} = \frac{\Gamma(n)}{(2i\pi)^n}\zeta(n)(1-p^{n-1}),$$

ce qui correspond à ce que l'on peut trouver dans le théorème 1.8. De même, le facteur d'Euler prédit par la conjecture dans le cas d'une forme modulaire coïncide avec celui de la proposition 1.15.

(v) Les fonctions $L_p(M(n)\otimes\psi,v)$ et $L_p(M\otimes\chi_{\mathrm{cycl}}^n\psi,v)$ sont reliées de manière simple mais ne coïncident pas. C'est dû au fait que le facteur Γ que l'on a introduit pour faire l'interpolation p-adique n'est pas le bon. En particulier, il dépend uniquement de la dimension de M_B^+ et pas de la filtration de Hodge sur M_{dR}. C'est une des raisons pour lesquelles on est forcé d'introduire des pseudo-distributions ; si n est suffisamment négatif, $\mu_{M(n),v}$ est une vraie distribution. Le chapitre II de [41] est consacré à la manière de remédier à ce problème. Remarquons tout de même que la formule donnant l'ordre du zéro de la fonction L p-adique est consistante avec celle donnant l'ordre de la fonction L complexe, compte tenu des pôles de la fonction Γ. Insistons aussi sur le fait que même avec les bons facteurs Γ, la distribution que l'on obtient n'est pas une mesure sauf dans des cas très particuliers.

(vi) On trouvera dans la section 2.11 une conjecture encore plus optimiste qui explique les normalisations (facteurs Γ et opérateur $\frac{1-p^{-1}\varphi^{-1}}{1-\varphi}$) et dans le chapitre consacré aux fonctions L d'Iwasawa la construction d'une fonction analytique p-adique ayant le même comportement (à une unité près) que celui conjecturé pour la fonction L p-adique attachée à M.

(vii) Si 1 ou p^{-1} est valeur propre de φ sur $M(n)_{\mathrm{dR},p}$, des zéros triviaux n'ayant pas toujours d'analogue complexe apparaissent (cf. [33] par exemple). Il y a eu récemment un certain nombre de résultats frappants concernant ces zéros triviaux, mais nous n'en dirons rien ; ils mériteraient un exposé séparé.

2.9. **Modules d'Iwasawa**

Soit $\Lambda = \mathbf{Z}_p[[\Gamma]]$ l'algèbre des mesures sur Γ à valeurs dans $\mathbf{Z}_p$. On l'appelle "algèbre d'Iwasawa" et on a $\mathscr{D}_0(\Gamma) = \mathbf{Q}_p \otimes \Lambda$. Si $\gamma \in \Gamma$, notons δ_γ la masse de Dirac en γ.

Si X est un Λ-module de torsion et de type fini, il existe $\mu_1,\ldots,\mu_r \in \Lambda$ et un morphisme de Λ-modules de X dans $\Lambda\mu_1\oplus\cdots\oplus\Lambda\mu_r$ dont le noyau et le conoyau sont finis. Le produit (de convolution) $\mu_1\cdots\mu_r$ qui ne dépend, à multiplication près par une unité de Λ, que de X sera noté μ_X et appelé "la mesure caractéristique de X".

Les idempotents e^+ et e^- introduits dans le paragraphe 1.2 permettent de décomposer Λ sous la forme $\Lambda^+ \times \Lambda^-$, où l'on a posé $\Lambda^+ = e^+\Lambda$ et $\Lambda^- = e^-\Lambda$. De manière plus générale, on peut décomposer tout Λ-module X comme produit d'un Λ^+-module X^+ et d'un Λ^--module X^-.

Si A est un $\mathbf{Z}_p$-module topologique complet (pas nécessairement de type fini ; en particulier les $\mathbf{Q}_p$-espaces vectoriels sont admis) muni d'une action continue de $\mathscr{G}_{\mathbf{Q},S}$, on munit l'espace $\Lambda \otimes A$ des mesures sur Γ à valeurs dans A de l'action diagonale de $\mathscr{G}_{\mathbf{Q},S}$: si $g \in \mathscr{G}_{\mathbf{Q},S}$, $a \in A$ et $\mu \in \Lambda$, on a $g(\mu \otimes a) = (\delta_{\overline{g}}\mu) \otimes g(a)$, où $\overline{g}$ désigne l'image de g dans Γ. Si $i \in \mathbf{N}$, on note alors $H^i_{\mathrm{Iw}}(A)$ le module $H^i(\mathscr{G}_{\mathbf{Q},S}, \Lambda \otimes A)$. Remarquons que, Γ étant commutatif, l'action de $\mathscr{G}_{\mathbf{Q},S}$ respecte la structure de Λ-module (i.e. $g(\lambda x) = \lambda g(x)$ si $g \in \mathscr{G}_{\mathbf{Q},S}$, $\lambda \in \Lambda$ et $x \in \Lambda \otimes A$) et donc que les $H^i_{\mathrm{Iw}}(A)$ ont une structure naturelle de Λ-module. On peut donner, grâce au lemme de Shapiro, une autre description de ces modules (cf. proposition 2.9).

On vérifie facilement que l'application qui à μ associe $\chi^k_{\mathrm{cycl}}\mu$ induit un isomorphisme $\mathscr{G}_{\mathbf{Q},S}$-équivariant de $\Lambda \otimes A$ sur $\Lambda \otimes A(k)$. D'autre part, si $\mu \in H^i_{\mathrm{Iw}}(A)$ et $\mu_{g_1,\dots,g_i}$ est un i-cocycle continu sur $\mathscr{G}_{\mathbf{Q},S}$ représentant μ, alors $\int_{\Gamma_n} \chi^k_{\mathrm{cycl}}\mu_{g_1,\dots,g_i}$ est un i-cocycle sur $\mathscr{G}_{\mathbf{Q}(\varepsilon_n),S}$ à valeurs dans $A(k)$ dont la classe $\int_{\Gamma_n} \chi^k_{\mathrm{cycl}}\mu$ dans $H^i(\mathscr{G}_{\mathbf{Q}(\varepsilon_n),S}, A(k))$ ne dépend que de μ et pas du choix du cocycle le représentant.

PROPOSITION 2.9. — *Si T est un $\mathbf{Z}_p$-module de type fini muni d'une action continue de $\mathscr{G}_{\mathbf{Q},S}$ et $i \in \mathbf{N}$, alors l'application qui à $\mu \in H^i_{\mathrm{Iw}}(T)$ associe $(\dots, \int_{\Gamma_n} \chi^k_{\mathrm{cycl}}\mu, \dots)$ induit un isomorphisme de $H^i_{\mathrm{Iw}}(T)$ sur la limite projective des $H^i(\mathscr{G}_{\mathbf{Q}(\varepsilon_n),S}, T(k))$ relativement aux applications de corestriction.*

Cette description est plus classique. Remarquons en particulier que l'on a, pour tout $k \in \mathbf{Z}$, une application $\mu \to \int_\Gamma \chi^k_{\mathrm{cycl}}\,\mu$ de $H^i_{\mathrm{Iw}}(T)$ dans $H^i(T(k))$, ce qui permet de voir $H^i_{\mathrm{Iw}}(T)$ comme une interpolation p-adique des modules $H^i(T(k))$. Le groupe $H^2_{\mathrm{Iw}}(\mathbf{Z}_p(1))^-$ se décrit classiquement comme la partie $-$ de la limite projective des p-sylow des groupes de classes d'idéaux des corps $\mathbf{Q}(\varepsilon_n)$ et la théorie de Kummer nous fournit une application de $\varprojlim \mathbf{Q}(\varepsilon_n)^*$ dans $H^1_{\mathrm{Iw}}(\mathbf{Z}_p(1))$ et si on note μ_{cycl} l'image du système des unités cyclotomiques (i.e. l'élément $(\varepsilon_n - 1)_{n\geq 1}$ de $\varprojlim \mathbf{Q}(\varepsilon_n)^*$), l'élément $\int_\Gamma \chi^{k-1}_{\mathrm{cycl}}\,\mu_{\mathrm{cycl}}$ de $H^1(\mathbf{Z}_p(k))$ est, si $k \geq 1$, "l'élément cyclotomique en K-théorie" introduit par Soulé [51].

2.10. **L'exponentielle de Perrin-Riou**

Soit $\ell \in S$. Soit $\mathscr{H}_{\mathbf{Q}_\ell} \subset \mathscr{G}_{\mathbf{Q}_\ell}$ le noyau du caractère cyclotomique de telle sorte que $\Gamma_{\mathbf{Q}_\ell} = \mathscr{G}_{\mathbf{Q}_\ell}/\mathscr{H}_{\mathbf{Q}_\ell}$ est le groupe de Galois $\mathrm{Gal}(\mathbf{Q}_\ell(\mu_{p^\infty})/\mathbf{Q}_\ell)$ et est un sous-groupe de Γ. Si A est un $\mathbf{Z}_p$-module topologique complet muni d'une action continue de $\mathscr{G}_{\mathbf{Q}_\ell}$ et si $i \in \mathbf{N}$, on note $H^i_{\mathrm{Iw}}(\mathbf{Q}_\ell, A)$ le Λ-module $H^i(\mathbf{Q}_\ell, \Lambda \otimes A)$.

PROPOSITION 2.10. — *Soient V une représentation p-adique de $\mathscr{G}_{\mathbf{Q}_\ell}$ et T un réseau de V stable par $\mathscr{G}_{\mathbf{Q}_\ell}$.*

(i) les modules $H^i_{\mathrm{Iw}}(\mathbf{Q}_\ell, T)$ *sont nuls si* $i \notin \{1,2\}$.

(ii) $H^2_{\mathrm{Iw}}(\mathbf{Q}_p, T)$ *est, à un groupe fini près, isomorphe à* $\mathrm{Ind}_{\Gamma_{\mathbf{Q}_\ell}}^{\Gamma} T(-1)^{\mathscr{H}_{\mathbf{Q}_\ell}}$ *en tant que* Λ*-module; en particulier il est de torsion.*

(iii) $H^1_{\mathrm{Iw}}(\mathbf{Q}_\ell, T)$ *est de torsion si* $\ell \neq p$ *et de rang* $\dim_{\mathbf{Q}_p} V$ *sur* Λ *si* $\ell = p$. *De plus le sous-module de torsion* $H^1_{\mathrm{Iw}}(\mathbf{Q}_\ell, T)_{\mathrm{tors}}$ *de* $H^1_{\mathrm{Iw}}(\mathbf{Q}_\ell, T)$ *est isomorphe à* $\mathrm{Ind}_{\Gamma_{\mathbf{Q}_\ell}}^{\Gamma} T^{\mathscr{H}_{\mathbf{Q}_\ell}}$.

La proposition ci-dessus dont on peut trouver la démonstration dans [39, 40] montre que, pour $\ell \neq p$, les modules $H^i_{\mathrm{Iw}}(\mathbf{Q}_\ell, T)$ sont relativement simples. Pour analyser plus en détail le module $H^1_{\mathrm{Iw}}(\mathbf{Q}_p, T)$, Perrin-Riou a introduit [40] une application exponentielle interpolant les applications exponentielles de Bloch-Kato pour les différents twists de V par les puissances du caractère cyclotomique. C'est dans la construction de cette application que l'opérateur $\frac{1-p^{-1}\varphi^{-1}}{1-\varphi}$ apparaît.

THÉORÈME 2.11. — *Si* V *est une représentation cristalline de* $\mathscr{G}_{\mathbf{Q}_p}$, *alors il existe une (unique) application*

$$\mathrm{Exp}_V : \mathrm{D}_{\mathrm{cris}}(V) \longrightarrow \tilde{\mathscr{P}}(\Gamma) \otimes_\Lambda H^1_{\mathrm{Iw}}(\mathbf{Q}_p, V)$$

telle que si k *est un entier assez grand et* $v \in \mathrm{D}_{\mathrm{cris}}(V)$, *alors*

$$\int_{\Gamma_K} \chi(x)^k \mathrm{Exp}_V(v) = \exp_{V(k)}\Big(\frac{1-p^{-1}\varphi^{-1}}{1-\varphi}\Big(\frac{\Gamma(k)}{(-t)^k}v\Big)\Big).$$

Le théorème suivant qui est l'une des formes équivalentes [42] de la loi de réciprocité conjecturée par Perrin-Riou (conjecture Réc(V) de [40]) montre que Exp_V interpole aussi les inverses des exponentielles duales $\exp^*_{V^*(1-k)}$ pour $k \ll 0$.

Rappelons que si $k \leq 0$, on a posé $\Gamma^*(k) = \lim_{s\to 0} s\Gamma(k+s) = \frac{(-1)^k}{k!}$

THÉORÈME 2.12. — *(i) Si* $k \ll 0$, *alors*

$$\int_\Gamma \chi_{\mathrm{cycl}}^k \mathrm{Exp}_V(v) = \big(\exp^*_{V^*(1-k)}\big)^{-1}\Big(\frac{1-p^{-1}\varphi^{-1}}{1-\varphi}\Big(\frac{\Gamma^*(k)}{(-t)^k}v\Big)\Big).$$

(ii) Si de plus, $\mathrm{D}_{\mathrm{cris}}(V)^{\varphi=1} = \mathrm{D}_{\mathrm{cris}}(V)^{\varphi=p^{-1}} = 0$, *alors*

$$\lim_{s\to 0} s\int_\Gamma \langle\chi_{\mathrm{cycl}}\rangle^s \mathrm{Exp}_V(v) = \exp_V\Big(\frac{1-p^{-1}\varphi^{-1}}{1-\varphi}v\Big)$$

et si $\frac{1-p^{-1}\varphi^{-1}}{1-\varphi}v \in \mathrm{D}^0_{\mathrm{cris}}(V)$, *alors* $\int_\Gamma \langle\chi_{\mathrm{cycl}}\rangle^s \mathrm{Exp}_V(v)$ *est analytique au voisinage de* $s = 0$ *et on a*

$$\frac{1-p^{-1}\varphi^{-1}}{1-\varphi}v = \exp_{V^*(1)}\Big(\int_\Gamma \mathrm{Exp}_V(v)\Big).$$

Ce théorème admet [40] comme corollaire le résultat suivant qui montre que si on étend Exp_V par linéarité à $\tilde{\mathscr{P}}(\Gamma) \otimes \mathrm{D}_{\mathrm{cris}}(V)$, alors Exp_V devient un isomorphisme de $\tilde{\mathscr{P}}(\Gamma) \otimes \mathrm{D}_{\mathrm{cris}}(V)$ sur $\tilde{\mathscr{P}}(\Gamma) \otimes H^1_{\mathrm{Iw}}(\mathbf{Q}_p, V)$. Son inverse sera noté Log_V.

PROPOSITION 2.13. — *Si $x_1, \dots, x_d$ est une base sur $\mathscr{D}_0(\Gamma) = \mathbf{Q}_p \otimes \Lambda$ de $H^1_{\mathrm{Iw}}(\mathbf{Q}_p, V)$ modulo torsion et si $v_1, \dots, v_d$ est une base de $\mathrm{D}_{\mathrm{cris}}(V)$ sur $\mathbf{Q}_p$, le déterminant de Exp_V dans ces bases est donné, à multiplication près par une unité de $\mathscr{D}_0(\Gamma)$, par la formule*

$$\mu_{H^2_{\mathrm{Iw}}(\mathbf{Q}_p,T)}\mu_{H^1_{\mathrm{Iw}}(\mathbf{Q}_p,T)_{\mathrm{tors}}}\frac{[\mathrm{Exp}_V(v_1),\dots\mathrm{Exp}_V(v_d)]}{[x_1,\dots,x_d]} = \prod_{i>0}\ell_i^{\dim \mathrm{D}^i_{\mathrm{cris}}(V)}\prod_{i\leq 0}\ell_i^{-d+\dim \mathrm{D}^i_{\mathrm{cris}}(V)}.$$

Si $V = \mathbf{Q}_p(1)$ et $u \in \varprojlim \mathscr{O}^*_{\mathbf{Q}_p(\varepsilon_n)}$, on note $\delta(u)$ son image dans $H^1_{\mathrm{Iw}}(\mathbf{Q}_p(1))$ donnée par la théorie de Kummer et $\mathrm{Log}_{\mathbf{Q}_p(1)}(\delta(u))$ est une distribution sur Γ à valeurs dans $t^{-1}\mathbf{Q}_p$ qui est celle obtenue en multipliant par t^{-1} la distribution sur Γ obtenue en tirant par $\chi_{\mathrm{cycl}}^{-1}$ la restriction à $\mathbf{Z}_p^*$ de la distribution λ_u fournie par Coleman, ce qui montre que Log_V est une vaste généralisation des séries de Coleman.

La démonstration initiale du théorème 2.11 se trouve dans [40]. Cette démonstration est décortiquée dans [14], où il est montré que l'exponentielle de Perrin-Riou s'exprime naturellement comme la composée de l'exponentielle de Bloch-Kato et d'une transformée de Fourier. Deux autres constructions se trouvent dans [5, 30]. Celle de Benois [5] repose sur la théorie des (φ,Γ)-modules introduite par Fontaine [18] pour classifier toutes les représentations p-adiques de $\mathscr{G}_{\mathbf{Q}_p}$ et celle de Kato, Kurihara et Tsuji [30] utilise la cohomologie syntomique et est assez semblable dans l'esprit à celle de Benois. D'autre part, [5, 14] et [30] contiennent aussi une démonstration de la loi de réciprocité de Perrin-Riou (sous une forme plus ou moins reconnaissable). En particulier, celle de [14] repose sur une construction directe de l'application Log_V qui s'étend d'ailleurs au cas des représentations de de Rham mais ne fournit pas de distribution tempérée dans ce cas (du moins pas de manière visible), ce qui rend problématique son utilisation pour la construction de fonctions L p-adiques. Elle utilise la proposition 2.6 alors que celles de [5] et [30] incluent la démonstration de cette proposition. Signalons aussi que l'application duale Exp^*_V de Exp_V se généralise [10] à toutes les représentations p-adiques de $\mathscr{G}_{\mathbf{Q}_p}$ grâce à la théorie des (φ,Γ)-modules et qu'une manière d'exprimer la loi de réciprocité de Perrin-Riou est de dire que Log_V et Exp^*_V coïncident (à normalisation près). D'autre part, le (ii) du théorème 2.12 peut se réécrire sous la forme suivante.

PROPOSITION 2.14. — *Si $\mathrm{D}_{\mathrm{cris}}(V)^{\varphi=1} = \mathrm{D}_{\mathrm{cris}}(V)^{\varphi=p^{-1}} = 0$, alors*

$$\frac{1-p^{-1}\varphi^{-1}}{1-\varphi}\Big(\int_\Gamma \mathrm{Log}_V(\mu)\Big) = \mathrm{exp}^*_{V^*(1)}\Big(\int_\Gamma \mu\Big).$$

Si de plus, $\int_\Gamma \mu \in H^1_f(\mathbf{Q}_p, V)$, alors

$$\mathrm{exp}_V\Big(\frac{1-p^{-1}\varphi^{-1}}{1-\varphi}\Big(\lim_{s\to 0} s^{-1}\int_\Gamma \langle\chi_{\mathrm{cycl}}\rangle^s \mathrm{Log}_V(\mu)\Big)\Big) = \int_\Gamma \mu.$$

Finalement, on a le résultat suivant [41, Chap. III] qui permet de calculer une partie de l'accouplement "hauteur p-adique" via les applications Exp_V et Log_V.

PROPOSITION 2.15. — *Si* $x \in H^1_f(V^*(1))$ *et* $\mu \in H^1_{\mathrm{Iw}}(V)$ *vérifie* $\int_\Gamma \mu \in H^1_f(V)$, *alors*

$$\langle x, \int_\Gamma \mu \rangle_{\mathrm{ht}} = \Big\langle s_{\mathrm{Hdg}}(\log_{V^*(1)}(y)), \frac{1-p^{-1}\varphi^{-1}}{1-\varphi}\Big(\lim_{s\to 0} s^{-1}\int_\Gamma \langle\chi_{\mathrm{cycl}}\rangle^s \mathrm{Log}_V(\mu)\Big)\Big\rangle_{\mathrm{dR}},$$

où $s_{\mathrm{Hdg}} : t_{V^*(1)} \to \mathrm{D}_{\mathrm{cris}}(V^*(1))$ *est le scindage de la filtration de Hodge de* $\mathrm{D}_{\mathrm{cris}}(V^*(1))$ *dual de celui de* $\mathrm{D}_{\mathrm{cris}}(V)$ *utilisé pour la définition de* $\langle\ ,\ \rangle_{\mathrm{ht}}$.

2.11. **Éléments spéciaux**

Soient M un motif défini sur $\mathbf{Q}$ et $k \in \mathbf{Z}$. On dit qu'un élément $\mu \in H^1_{\mathrm{Iw}}(M_{B,p})$ est $M(k)$-motivique si $\int_{\Gamma_n} \chi^k_{\mathrm{cycl}}\mu$ appartient au sous-$\mathbf{Q}$-espace vectoriel $H^1(\mathbf{Q}(\varepsilon_n), M(k))$ de $H^1(\mathscr{G}_{\mathbf{Q}(\varepsilon_n),S}, M_{B,p}(k))$ quel que soit $n \in \mathbf{N}$. Un élément $\mu \in H^1_{\mathrm{Iw}}(M_{B,p})$ est dit motivique s'il existe $k \in \mathbf{Z}$ tel que μ soit $M(k)$-motivique et fortement motivique s'il existe $k_0 \in \mathbf{Z}$ tel que μ soit $M(k)$-motivique quel que soit $k \geq k_0$. Le seul exemple d'élément fortement motivique dont on dispose à l'heure actuelle est celui des unités cyclotomiques qui, quand on les tord, donnent les éléments de Beilinson [7] (la démonstration consiste à comparer les éléments de Soulé et ceux de Beilinson et ne semble pas être de tout repos [8, 26, 25]). D'autre part, si f est une forme modulaire (pour un sous-groupe de congruence de $\mathrm{SL}_2(\mathbf{Z})$), Kato a construit[2] un élément motivique μ_{Kato} de $H^1_{\mathrm{Iw}}(M(f)_{B,p})$, où $M(f)$ est le motif associé à f. Il faut quand même bien voir que ces deux exemples sont un peu particuliers car on est en rang 1 (i.e $d^-(M) = 1$).

CONJECTURE 2.16. — *Soit* M *un motif défini sur* $\mathbf{Q}$ *vérifiant* $d^- \neq 0$. *Il existe une base* ω_{dR} *une base de* M_{dR} *(sur* $\mathbf{Q}$*) et une famille* $\mu = (\mu_1, \ldots, \mu_{d^-})$ *d'éléments motiviques de* $H^1_{\mathrm{Iw}}(M_{B,p})$ *tels que l'on ait*

$$L_p(M, v, s) = \Big(\prod_{i>0}(s+i)^{\dim M^i_{\mathrm{dR}}} \prod_{i\leq 0}(s+i)^{\dim M^i_{\mathrm{dR}}-d}\Big)\frac{\big[\int_\Gamma \langle\chi_{\mathrm{cycl}}\rangle^s \mathrm{Log}_{M_{B,p}}(\mu)\oplus v^+\big]}{[\omega_{\mathrm{dR}}]}$$

quelle que soit la famille $v = (v_1, \ldots, v_d)$ *de* $M_{\mathrm{dR},p}$.

On peut supprimer les termes parasites dans le membre de droite en introduisant des facteurs Γ convenables dans la définition de la fonction L p-adique. Cette conjecture est particulièrement optimiste et les seuls résultats que l'on ait à son sujet concernent d'une part, le motif $\mathbf{Q}(1)$ pour lequel on retombe sur la construction de Coleman de la fonction zêta de Kubota-Leopoldt et d'autre part, le motif d'une forme modulaire pour lequel Kato a vérifié que l'on retombe bien sur la fonction L p-adique de la forme modulaire à partir de μ_{Kato}. Remarquons qu'une telle construction de la fonction L p-adique d'un motif fournit, via la proposition 2.14, des renseignements extrêmement précis sur les valeurs aux entiers de la fonction L p-adique.

[2] La construction de Kato et les résultats qui en découlent n'ont malheureusement pas encore été publiés (voir toutefois [48]) ce qui est dommage car ils constituent, avec les travaux de Wiles sur la conjecture de Taniyama-Weil, le plus bel ensemble de résultats de la dernière décennie en théorie des nombres.

3. FONCTIONS L D'IWASAWA

3.1. La conjecture de Leopoldt faible

Rappelons que la conjecture de Leopoldt pour un corps de nombres F est équivalente à la nullité de $H^2(G_{F,S}, \mathbf{Q}_p/\mathbf{Z}_p)$ (c'est une des nombreuses reformulations de la conjecture de Leopoldt). La conjecture ci-dessous qui apparaît dans les travaux de Schneider [49] et Greenberg [22] en est une généralisation affaiblie et est connue sous le nom de conjecture de Leopoldt faible (pour la représentation V).

CONJECTURE 3.1. — *Si V est une représentation p-adique de $\mathscr{G}_{\mathbf{Q},S}$ et T un réseau de V stable par $\mathscr{G}_{\mathbf{Q},S}$, alors il existe $k \in \mathbf{Z}$ tel que $H^2(V(k)/T(k)) = 0$.*

La conjecture de Leopoldt faible est une conséquence de la plupart des conjectures raisonnables du sujet et on a un certain nombre de critères numériques permettant de la vérifier sur des cas particuliers. Par exemple, elle découle de la formule donnant l'ordre du zéro de la fonction L complexe dans la conjecture 2.1 et de la conjecture 2.3 ; on a donc intérêt à ce qu'elle soit vraie sinon le bel édifice présenté dans les sections précédentes s'écroule inexorablement. Elle sera absolument fondamentale dans la suite. Elle admet en particulier comme conséquence le résultat suivant [41, Chap. I].

PROPOSITION 3.2. — *Si V vérifie la conjecture de Leopoldt faible, alors*
(i) $H^2_{\mathrm{Iw},S}(V)$ est de torsion,
(ii) $H^1_{\mathrm{Iw},S}(V)^+$ est de rang d^- sur Λ^+ et $H^1_{\mathrm{Iw}}(T)^-$ est de rang d^+ sur Λ^-.

3.2. La fonction L d'Iwasawa d'une représentation p-adique

Dans tout le reste du texte, on suppose que V et $V^*(1)$ vérifient la conjecture de Leopoldt faible et donc, en particulier, que $H^1_{\mathrm{Iw}}(T)^+$ est de rang d^- sur Λ^+ et $H^1_{\mathrm{Iw}}(T)^-$ est de rang d^+ sur Λ^-. Soit $x = (x_1, \dots, x_d)$ une famille d'éléments de $H^1_{\mathrm{Iw}}(\mathbf{Q}, T)$ vérifiant les trois conditions suivantes
(i) $x_1, \dots, x_{d^-} \in H^1_{\mathrm{Iw}}(T)^+$
(ii) $x_{d^-+1}, \dots, x_d \in H^1_{\mathrm{Iw}}(T)^-$
(iii) $H(x) = H^1_{\mathrm{Iw}}(T)/(\Lambda x_1 + \cdots + \Lambda x_d)$ est un Λ-module de torsion.
Si $v = (v_1, \dots, v_d)$ est une famille de d vecteurs de $\mathrm{D}_{\mathrm{cris}}(V)$, on définit la famille $y_v = (y_{v,1}, \dots, y_{v,d})$ d'éléments de $\tilde{\mathscr{P}}(\Gamma) \otimes H^1_{\mathrm{Iw}}(\mathbf{Q}_p, T)$ par la formule

$$y_{v,i} = \begin{cases} x_i + e^-(\mathrm{Exp}_V(v_{d^++i})) & \text{si } 1 \leq i \leq d^- \\ x_i + e^+(\mathrm{Exp}_V(v_{i-d^-})) & \text{si } d^- + 1 \leq i \leq d \end{cases}.$$

Comme tensoriser par $\tilde{\mathscr{P}}(\Gamma)$ tue le sous-Λ-module de torsion de $H^1_{\mathrm{Iw}}(\mathbf{Q}_p, T)$, on peut trouver une famille $z = (z_1, \dots, z_d)$ d'éléments de $H^1_{\mathrm{Iw}}(\mathbf{Q}_p, T)$ formant une base de

$\tilde{\mathscr{D}}(\Gamma) \otimes H^1_{\mathrm{Iw}}(\mathbf{Q}_p, T)$ sur $\tilde{\mathscr{D}}(\Gamma)$ et le Λ-module $H(z) = H^1_{\mathrm{Iw}}(\mathbf{Q}_p, T)/(\Lambda z_1 + \cdots \Lambda z_d)$ est alors de torsion. L'élément

$$\mu^1_{T,v} = \frac{[y_v]}{[z]} \frac{\mu_{H(x)}}{\mu_{H(z)}}$$

de $\tilde{\mathscr{D}}(\Gamma)$ ne dépend pas des choix des x_i ou des z_i. Soient finalement

$$\mu^2_T = \mu_{H^2_{\mathrm{Iw}}(T)} \prod_{\ell \in S} (\mu_{H^2_{\mathrm{Iw}}(\mathbf{Q}_\ell, T)})^{-1} \quad \text{et} \quad \mu_{T,v} = \mu^1_{T,v} \mu^2_T \in \tilde{\mathscr{D}}(\Gamma).$$

On définit alors la fonction L d'Iwasawa attachée à V par les formules

$$L_{\mathrm{Iw}}(T \otimes \psi, v) = \int_\Gamma \psi \, \mu_{T,v} \quad \text{et} \quad L_{\mathrm{Iw}}(T, v, s) = \int_\Gamma \langle \chi_{\mathrm{cycl}} \rangle^s \mu_{T,v}.$$

Cette fonction L d'Iwasawa n'est bien définie qu'à multiplication près par une unité dans l'algèbre d'Iwasawa et dépend (un peu) de T et (beaucoup) de v.

3.3. La conjecture de Bloch-Kato à une unité près

On suppose dorénavant pour simplifier $p \geq 3$. Si M est un motif ayant bonne réduction en p, on dispose conjecturalement de deux fonctions-L p-adiques, à savoir la fonction $L_p(M \otimes \psi, v)$ construite en interpolant les valeurs de la fonction L complexe attachée à M et la fonction L d'Iwasawa $L_{\mathrm{Iw}}(M \otimes \psi, v)$ attachée à la représentation p-adique $M_{B,p}$.

CONJECTURE 3.3. — *Il existe une unité α de $\mathscr{D}_0(\Gamma)$ telle que, quels que soient le caractère continu ψ et la famille v, on ait*

$$L_p(M \otimes \psi, v) = \Big(\int_\Gamma \psi \, \alpha \Big) L_{\mathrm{Iw}}(T \otimes \psi, v).$$

Cette conjecture est connue sous le nom de "conjecture principale". Elle est démontrée en particulier dans le cas des motifs attachés aux caractères de Dirichlet grâce aux travaux de Mazur et Wiles [34] ou de Kolyvagin et Rubin [31, 44] et dans le cas d'une forme modulaire, les résultats de Kato auxquels il a été fait allusion permettent de démontrer une relation de divisibilité entre les deux fonctions L. Le théorème ci-dessous est le résultat général le plus encourageant en direction de cette conjecture : il montre que la fonction L d'Iwasawa vérifie la conjecture de Bloch-Kato (à une unité p-adique près).

Choisissons une base ω_{tg} de t_V. Si v est une famille de d vecteurs de $\mathrm{D}_{\mathrm{cris}}(V)$, définissons la période p-adique de T comme le discriminant

$$\Omega_p(T, v) = \big\langle H^1_f(T^*(1)) \oplus \omega^*_{\mathrm{tg}}, H^1_f(T) \oplus v^+ \big\rangle_{\mathrm{ht}}$$

relativement à l'accouplement "hauteur p-adique".

THÉORÈME 3.4 (Perrin-Riou). — *Si $\mathrm{D}_{\mathrm{cris}}(V)^{\varphi=1} = \mathrm{D}_{\mathrm{cris}}(V)^{\varphi=p^{-1}} = 0$, alors $L_{\mathrm{Iw}}(T, v, s)$ a, en $s = 0$, un zéro d'ordre au moins $r_p(V) = \dim_{\mathbf{Q}_p} H^1_f(\mathbf{Q}, V^*(1)) - d^+(V)$ et*

$$\lim_{s \to 0} s^{-r_p(V)} L_{\mathrm{Iw}}(T, v, s) \sim_p E_p(V, 1) |Ш(T^*(1))| \, \mathrm{Tam}^0_{\omega_{\mathrm{tg}}}(T) \, \Omega_p\Big(T, \frac{1 - p^{-1}\varphi^{-1}}{1 - \varphi} v\Big).$$

La démonstration de ce théorème utilise un certain nombre de suites exactes fournies par la théorie d'Iwasawa ou la dualité de Poitou-Tate (voir plus loin). On sépare l'étude de la fonction $L_{\mathrm{Iw}}(T,v,s)$ en celles des fonctions $\int_\Gamma \langle\chi_{\mathrm{cycl}}\rangle^s \mu^1_{T,v}$ et $\int_\Gamma \langle\chi_{\mathrm{cycl}}\rangle^s \mu^2_T$ qui ont un sens prises séparément (on montre, en utilisant des suites exactes de localisation [41, Chap. I], que $H^1_{\mathrm{Iw}}(T)$ ne dépend pas du choix de l'ensemble fini S qui intervient dans sa définition et que le noyau de l'application de $H^2_{\mathrm{Iw}}(T)$ dans $\oplus_{\ell\in S} H^2_{\mathrm{Iw}}(\mathbf{Q}_\ell,T)$ n'en dépend pas non plus). Les formules obtenues pour le comportement en $s=0$ de ces fonctions prises séparément sont assez compliquées et il y a un certain nombre de simplifications qui apparaissent quand on les réunit. De manière précise, soient $H^1_f(V^*(1))_0$, $H^1_f(T^*(1))_0$ et $H^1_f(T^\vee(1))_0$ respectivement les sous-groupes des éléments de $H^1_f(V^*(1))$, $H^1_f(T^*(1))$ et $H^1_f(T^\vee(1))$ dont la restriction à $\mathscr{G}_{\mathbf{Q}_p}$ est triviale (le 0 en indice signifie "0 en p"). Introduisons aussi les images respectives $H^1(V)_u$ et $H^1(T)_u$ de $H^1_{\mathrm{Iw}}(V)_\Gamma$ et $H^1_{\mathrm{Iw}}(T)_\Gamma$ dans $H^1(V)$ et $H^1(T)$ et soient $H^1_f(V)_u = H^1(V)_u \cap H^1_f(V)$ et $H^1_f(T)_u = H^1(T)_u \cap H^1_f(T)$ (le u en indice signifie "universel" comme dans "norme universelle"). Si on introduit alors les constantes

$$C_p = [H^1(\mathbf{Q}_p,T) : H^1_{\mathrm{Iw}}(\mathbf{Q}_p,T)_\Gamma],$$
$$C_u = \Big[\frac{H^1_{f,\{p\}}(T)}{H^1(T)_u} : \frac{H^1_f(T)}{H^1_f(T)_u}\Big],$$
$$C_{f,0} = \langle (H^1_f(T^\vee(1))_0)^\vee, H^1_f(T^*(1))_0\rangle,$$
$$C_{f,1} = \Big\langle \Big(\frac{H^1_f(T^\vee(1))}{H^1_f(T^\vee(1))_0}\Big)^\vee, \frac{H^1_f(T^*(1))}{H^1_f(T^*(1))_0}\Big\rangle,$$

on a les résultats suivants.

PROPOSITION 3.5. — *La fonction $s\to \int_\Gamma \langle\chi_{\mathrm{cycl}}\rangle^s \mu^2_T$ a un zéro en $s=0$ d'ordre au moins égal à $r_{p,2}(V) = \dim_{\mathbf{Q}_p} H^1_f(V^*(1))_0$ et, si l'injection naturelle de $H^1_f(V)/H^1_f(V)_u$ dans $H^1_{f,\{p\}}(V)/H^1(V)_u$ est une bijection, on a*

$$\lim_{s\to 0} s^{-r_{p,2}(V)} \int_\Gamma \langle\chi_{\mathrm{cycl}}\rangle^s \mu^2_T \sim_p C_p C_u^{-1} C_{f,0}^{-1} \frac{\prod_{\ell\in S-\{p\}} \mathrm{Tam}^0_\ell(T)}{|H^0(\mathbf{Q},T^\vee(1))|} \Big\langle H^1_f(T^*(1))_0, \frac{H^1_f(T)}{H^1_f(T)_u}\Big\rangle_{\mathrm{ht}}.$$

L'ingrédient principal pour démontrer cette proposition est une réinterprétation de l'accouplement "hauteur p-adique" en termes de théorie d'Iwasawa (cf. proposition 3.9 ci-après).

PROPOSITION 3.6. — *(i) La fonction $s\to \int_\Gamma \langle\chi_{\mathrm{cycl}}\rangle^s \mu^1_{T,v}$ a, en $s=0$, un zéro d'ordre au moins $r_{p,1}(V) = -d^+ + \dim H^1_f(V^*(1)) - \dim H^1_f(V^*(1))_0$, cette inégalité étant stricte si l'injection naturelle de $H^1_f(V)/H^1_f(V)_u$ dans $H^1_{f,\{p\}}(V)/H^1(V)_u$ n'est pas une bijection.*

(ii) Si l'injection naturelle de $H^1_f(V)/H^1_f(V)_u$ dans $H^1_{f,\{p\}}(V)/H^1(V)_u$ est une bijection, alors $\lim_{s\to 0} s^{-r_{p,1}(V)} \int_\Gamma \langle\chi_{\mathrm{cycl}}\rangle^s \mu^1_{T,v}$ est égal (à une unité p-adique près) à

$$C_p^{-1} C_u C_{f,1}^{-1} E_p(V,1) \mathrm{Tam}^0_{p,\omega_{\mathrm{tg}}}(T) \Big\langle \omega^*_{\mathrm{tg}} \oplus \frac{H^1_f(T^*(1))}{H^1_f(T^*(1))_0}, H^1_f(T)_u \oplus \frac{1-p^{-1}\varphi^{-1}}{1-\varphi} v^+\Big\rangle_{\mathrm{ht}}.$$

L'ingrédient principal dans la démonstration de cette proposition est la loi de réciprocité explicite et plus exactement sa conséquence énoncée dans la proposition 2.14.

Modulo le fait que $H^1_f(V^*(1))_0$ est orthogonal à $H^1_f(V)_u \oplus \mathrm{D}_{\mathrm{cris}}(V)$ (cf. proposition 3.9), ce qui est d'ailleurs implicite dans la formulation des deux propositions ci-dessus, le théorème est une conséquence immédiate de ces propositions une fois que l'on a remarqué que l'on a

$$C_{f,0}C_{f,1} = \left\langle H^1_f(T^\vee(1))^\vee, H^1_f(T^*(1)) \right\rangle = |Ш(T^*(1))|^{-1}|H^0(T^\vee(1))|^{-1},$$

comme on le voit en utilisant les deux suites exactes

$$0 \longrightarrow Ш(T^*(1))^\vee \longrightarrow H^1_f(T^\vee(1))^\vee \longrightarrow H^1_f(T^*(1))^* \longrightarrow 0,$$
$$0 \longrightarrow H^0(T^\vee(1)) \longrightarrow H^1_f(T^*(1)) \longrightarrow (H^1_f(T^*(1))^*)^* \longrightarrow 0.$$

3.4. **Outils de théorie d'Iwasawa**

Comme on a supposé $p \neq 2$, le groupe Γ est procyclique. Soit γ un de ses générateurs topologiques. Si X est un Λ-module, on note X^Γ le noyau de $\delta_\gamma - \delta_1$ et X_Γ le module $X/(\delta_\gamma - \delta_1)X$. Si $0 \to A \to B \to C \to 0$ est une suite exacte de Λ-module, le lemme du serpent fournit une suite exacte

$$0 \to A^\Gamma \to B^\Gamma \to C^\Gamma \to A_\Gamma \to B_\Gamma \to C_\Gamma \to 0$$

de $\mathbf{Z}_p$-modules (remarquons tout de même que l'application de connexion $C^\Gamma \to A_\Gamma$ dépend du choix de γ).

La proposition suivante qui découle très facilement du théorème de structure des Λ-modules de torsion (et de type fini) est extrêmement utile pour l'étude du comportement des fonctions L sortant de la théorie d'Iwasawa.

PROPOSITION 3.7. — *Si X est un Λ-module de torsion (et de type fini), alors $s \to \int_\Gamma \langle\chi_{\mathrm{cycl}}\rangle^s \mu_X$ a un zéro d'ordre au moins $r_0(X) = \dim_{\mathbf{Q}_p} \mathbf{Q}_p \otimes X^\Gamma = \dim_{\mathbf{Q}_p} \mathbf{Q}_p \otimes X_\Gamma$ en $s = 0$, avec égalité si et seulement si l'application naturelle (induite par l'identité sur X) de $\mathbf{Q}_p \otimes X^\Gamma$ dans $\mathbf{Q}_p \otimes X_\Gamma$ est un isomorphisme auquel cas on a*

$$\lim_{s\to 0} s^{-r_0(X)} \int_\Gamma \langle\chi_{\mathrm{cycl}}\rangle^s \mu_X \sim_p (\log \chi(\gamma))^{r_0(X)} [X_\Gamma : X^\Gamma],$$

où $[X_\Gamma : X^\Gamma]$ est l'indice généralisé de X^Γ dans X_Γ c'est-à-dire le quotient du cardinal du conoyau par celui du noyau de l'application de X^Γ dans X_Γ induite par l'identité sur X.

Finalement, si A est un $\mathbf{Z}_p$-module, on dispose de suites exactes

$$0 \to \Lambda \otimes T \to \Lambda \otimes T \to T \to 0 \quad \text{et} \quad 0 \to T \to \mathscr{C}^0(\Gamma, T) \to \mathscr{C}^0(\Gamma, T) \to 0,$$

l'application de $\Lambda \otimes T$ dans $\Lambda \otimes T$ étant la mutiplication par $\delta_\gamma - \delta_1$, celle de $\Lambda \otimes T$ dans T étant l'intégration sur Γ, celle de T dans $\mathscr{C}^0(\Gamma, T)$ étant celle qui à v associe la fonction constante de valeur v et celle de $\mathscr{C}^0(\Gamma, T)$ dans $\mathscr{C}^0(\Gamma, T)$ étant l'application qui,

à une fonction continue $x \to f(x)$ sur Γ associe la fonction $x \to f(\gamma x) - f(x)$; c'est la transposée de la multiplication par $\delta_\gamma - \delta_1$ si l'on voit les fonctions continues sur Γ comme le dual de Λ. Si A est muni d'une action continue de $\mathscr{G}_{\mathbf{Q},S}$, les suites exactes ci-dessus sont $\mathscr{G}_{\mathbf{Q},S}$-équivariantes ce qui nous fournit des suites exactes longues de cohomologie reliant les $H^i_{\mathrm{Iw}}(A)$ aux $H^i(A)$. Les suites exactes courtes que l'on tire de ces suites exactes longues sont d'ailleurs équivalentes à celles que l'on obtient via les suites exactes d'inflation-restriction si on utilise la description des modules d'Iwasawa donnée dans la proposition 2.9.

3.5. **Suites exactes de Poitou-Tate**

Si A est un module topologique discret, les groupes de cohomologie $H^i(A)$ et $H^i(A^\vee(1))$ sont reliés par la suite exacte suivante [35] dite "suite exacte de Poitou-Tate" :

$$\begin{array}{ccccccc}
0 & \longrightarrow & H^0(A^\vee(1)) & \longrightarrow & \oplus_{\ell\in S} H^0(\mathbf{Q}_\ell, A^\vee(1)) & \longrightarrow & H^2(A)^\vee \\
 & & & \searrow & & \nearrow & \\
 & & H^1(A^\vee(1)) & \longrightarrow & \oplus_{\ell\in S} H^1(\mathbf{Q}_\ell, A^\vee(1)) & \longrightarrow & H^1(A)^\vee \\
 & & & \searrow & & \nearrow & \\
 & & H^2(A^\vee(1)) & \longrightarrow & \oplus_{\ell\in S} H^2(\mathbf{Q}_\ell, A^\vee(1)) & \longrightarrow & H^0(A)^\vee \longrightarrow 0.
\end{array}$$

Si V est une représentation p-adique de $\mathscr{G}_{\mathbf{Q},S}$ et T est un réseau de V, on peut appliquer ce qui précède à $T^\vee(1)$ ou à $\mathscr{C}^0(\Gamma, T^\vee(1))$ espace des fonctions continues de Γ dans $T^\vee(1)$. Le dual de Pontryagin de $\mathscr{C}^0(\Gamma, T^\vee(1))$ est $\Lambda \otimes T(-1)$ et le lemme de Shapiro permet de montrer que l'on a $H^i(\mathscr{G}_{\mathbf{Q},S}, \mathscr{C}^0(\Gamma, T^\vee(1))) = H^i(\mathscr{G}_{\mathbf{Q}(\mu_{p^\infty}),S}, T^\vee(1))$. Si $i \in \mathbf{N}$, on note $X^i_{\mathrm{Iw}}(T)$ le dual de Pontryagin de $H^i(\mathscr{G}_{\mathbf{Q}(\mu_{p^\infty}),S}, T^\vee(1))$. La conjecture de Leopoldt faible pour $V^*(1)$ implique que $X^2_{\mathrm{Iw}}(T) = 0$ car il est de torsion et d'autre part il s'injecte dans un Λ-module libre [41, Chap. I]. La suite exacte de Poitou-Tate pour $\mathscr{C}^0(\Gamma, T^\vee(1))$ devient donc

$$0 \to H^1_{\mathrm{Iw}}(T) \to \oplus_{\ell\in S} H^1_{\mathrm{Iw}}(\mathbf{Q}_\ell, T) \to X^1_{\mathrm{Iw}}(T) \to H^2_{\mathrm{Iw}}(T) \to \oplus_{\ell\in S} H^2_{\mathrm{Iw}}(\mathbf{Q}_\ell, T) \to X^0_{\mathrm{Iw}}(T) \to 0.$$

On peut modifier la suite exacte de Poitou-Tate pour $T^\vee(1)$ en remplaçant le terme du milieu $\oplus_{\ell\in S} H^1(\mathbf{Q}_\ell, T)$ par un de ses sous-groupes H. Il faut alors remplacer $H^1(T)$ par l'image réciproque de H dans $H^1(T)$ et on tronque la suite obtenue en considérant le sous-groupe de $H^1(T^\vee(1))$ image réciproque du dual de T dans $\oplus_{\ell\in S} H^1(\mathbf{Q}_\ell, T^\vee(1))$. On obtient en particulier les suites exactes

$$\cdots \to H^1_f(T) \to \oplus_{\ell\in S} H^1_f(\mathbf{Q}_\ell, T) \to H^1(T^\vee(1))^\vee \to H^1_f(T^\vee(1))^\vee \to 0,$$

$$\cdots \to H^1_{f,\{p\}}(T) \to H^1(\mathbf{Q}_p, T) \oplus \left(\oplus_{\ell\in S-\{p\}} H^1_f(\mathbf{Q}_\ell, T)\right)$$

$$\downarrow$$

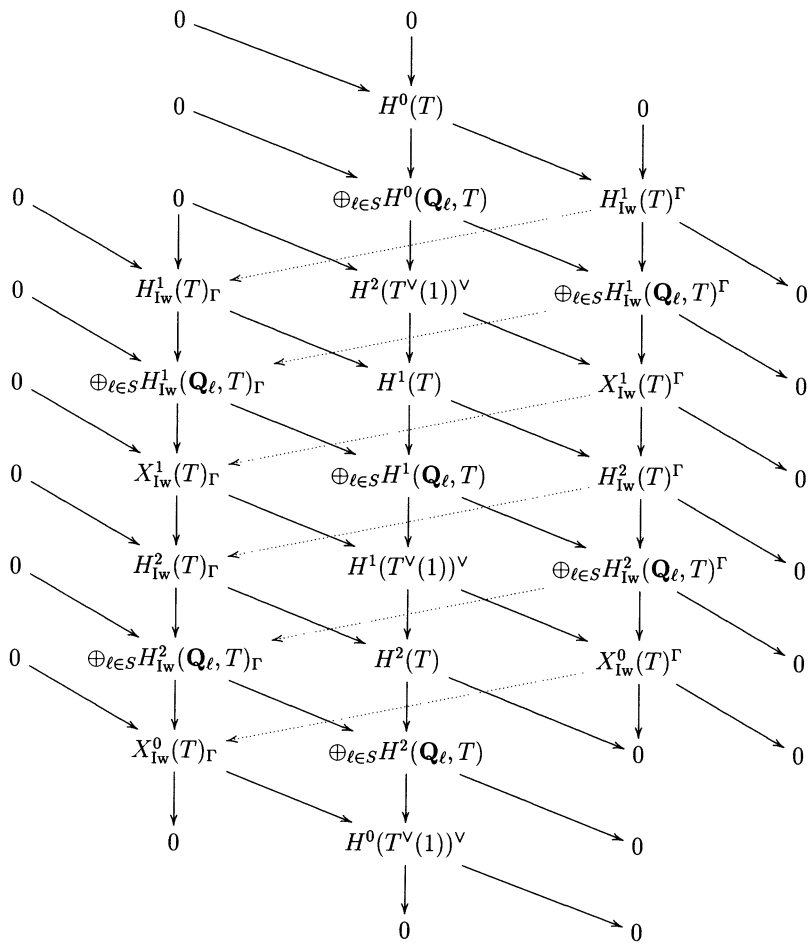
$$H^1(T^\vee(1))^\vee \to (H^1_f(T^\vee(1))_0)^\vee \to 0,$$

ainsi que la suite exacte

$$0 \to H^1_f(T) \to H^1_{f,\{p\}}(T) \to H^1(\mathbf{Q}_p, T)/H^1_f(\mathbf{Q}_p, T) \to H^1_f(T^\vee(1))^\vee \to (H^1_f(T^\vee(1))_0)^\vee \to 0$$

qui s'obtient en comparant les deux suites exactes précédentes.

3.6. **Un diagramme récapitulatif**



Le diagramme précédent résume ce que nous avons dit dans les deux précédentes sections. La colonne du milieu est la suite exacte de Poitou-Tate associée à $T^\vee(1)$ et les deux autres colonnes proviennent de la suite exacte de Poitou-Tate associée à $\mathscr{C}^0(\Gamma, T^\vee(1))$; elles ne sont pas exactes mais forment des complexes de $\mathbf{Z}_p$-modules. Les diagonales (en

trait plein) proviennent des suites exactes longues de cohomologie associées à la suite exacte $0 \to \Lambda \otimes T \to \Lambda \otimes T \to T \to 0$ et à sa duale. Elles sont exactes et montrent que la théorie d'Iwasawa sert à casser en deux les groupes $H^i(T)$. Finalement, la proposition 3.7 montre que pour démontrer le théorème 3.4, il s'agit d'étudier les flèches en pointillé.

3.7. Étude de $\mu^1_{T,v}$

Soient $a = \dim_{\mathbf{Q}_p} H^1_f(V)_u$ et $b = \dim_{\mathbf{Q}_p} H^1_f(\mathbf{Q}_p, V) = \dim_{\mathbf{Q}_p} t_V$. Choisissons les familles $x = (x_1, \dots, x_d)$ et $z = (z_1, \dots, z_d)$ utilisées pour la définition de $\mu^1_{T,v}$ de telle sorte que les conditions suivantes soient vérifiées :

(i) $\int_\Gamma x_1, \dots, \int_\Gamma x_{d^-}$ est une base de $H^1(V)_u$ sur $\mathbf{Q}_p$,
(ii) $\int_\Gamma x_1, \dots, \int_\Gamma x_a$ est une base de $H^1_f(V)_u$ sur $\mathbf{Q}_p$.
(iii) $\int_\Gamma z_1, \dots, \int_\Gamma z_d$ est une base de $H^1(\mathbf{Q}_p, V)$ sur $\mathbf{Q}_p$,
(iv) $\int_\Gamma z_1, \dots, \int_\Gamma z_b$ est une base de $H^1_f(\mathbf{Q}_p, V)$ sur $\mathbf{Q}_p$.

Maintenant, on a

$$\int_\Gamma \langle \chi_{\mathrm{cycl}} \rangle^s \frac{[y_v]}{[z]} = \int_\Gamma \langle \chi_{\mathrm{cycl}} \rangle^s \frac{[\mathrm{Log}_V(y_v)]}{[\mathrm{Log}_V(z)]} = \frac{\left[\int_\Gamma \langle \chi_{\mathrm{cycl}} \rangle^s \mathrm{Log}_V(y_v)\right]}{\left[\int_\Gamma \langle \chi_{\mathrm{cycl}} \rangle^s \mathrm{Log}_V(z)\right]}.$$

D'autre part, si on revient à la définition de $y_{v,i}$ et que l'on utilise le fait $x_i \in H^1_{\mathrm{Iw}}(T)^-$ si $i \geq d^- + 1$, on obtient

$$\int_\Gamma \langle \chi_{\mathrm{cycl}} \rangle^s \mathrm{Log}_V(y_{v,i}) = \begin{cases} \int_\Gamma \langle \chi_{\mathrm{cycl}} \rangle^s \mathrm{Log}_V(x_i) & \text{si } i \leq d^- \\ v_{i-d^-} & \text{si } i \geq d^- + 1 \end{cases}.$$

On peut alors utiliser la proposition 2.14 pour montrer que $\int_\Gamma \langle \chi_{\mathrm{cycl}} \rangle^s \frac{[y_v]}{[z]}$ a un zéro d'ordre au moins $a - b$ en $s = 0$ et obtenir une formule explicite pour $\lim_{s \to 0} s^{b-a} \int_\Gamma \langle \chi_{\mathrm{cycl}} \rangle^s \frac{[y_v]}{[z]}$ comme quotient de deux déterminants de familles de vecteurs de $\mathrm{D}_{\mathrm{cris}}(V)$. Pour passer de cette formule à celle de la proposition, on passe d'un quotient de déterminants dans $\mathrm{D}_{\mathrm{cris}}(V)$ à un discrimant utilisant l'accouplement $\langle\ ,\ \rangle_{\mathrm{dR}}$ sur $\mathrm{D}_{\mathrm{cris}}(V^*(1)) \times \mathrm{D}_{\mathrm{cris}}(V)$ en utilisant la formule

$$\langle \omega^*_{\mathrm{tg}}, \log_V(H^1_f(\mathbf{Q}_p, T)) \rangle^{-1} = E_p(V, 1) \mathrm{Tam}^0_{p, \omega_{\mathrm{tg}}}(T).$$

La suite exacte

$$0 \longrightarrow \frac{H^1_{f,\{p\}}(T)}{H^1_f(T)} \longrightarrow \frac{H^1(\mathbf{Q}_p, T)}{H^1_f(\mathbf{Q}_p, T)} \longrightarrow \left(\frac{H^1_f(T^\vee(1))}{H^1_f(T^\vee(1))_0}\right)^\vee \longrightarrow 0$$

qui est une suite exacte de Poitou-Tate un peu modifiée permet alors de casser ce discriminant en deux morceaux et de montrer que $a - b \geq \dim(H^1_f(V^*(1))/H^1_f(V^*(1))_0) - d^+$ avec égalité si et seulement si l'injection de $H^1_f(V)/H^1_f(V)_u$ dans $H^1_{f,\{p\}}(V)/H^1(V)_u$ est une bijection. L'un de ces morceaux s'évalue en utilisant la suite exacte

$$0 \longrightarrow \frac{H^1(T)_u}{H^1_f(T)_u} \longrightarrow \frac{H^1_{f,\{p\}}(T)}{H^1_f(T)} \longrightarrow \left(\frac{H^1_{f,\{p\}}(T)}{H^1(T)_u}\right) / \left(\frac{H^1_f(T)}{H^1_f(T)_u}\right) \longrightarrow 0$$

que l'on obtient via le lemme du serpent et l'autre est relié au discrimant de $\langle\ ,\ \rangle_{\mathrm{ht}}$ grâce à la proposition 2.15.

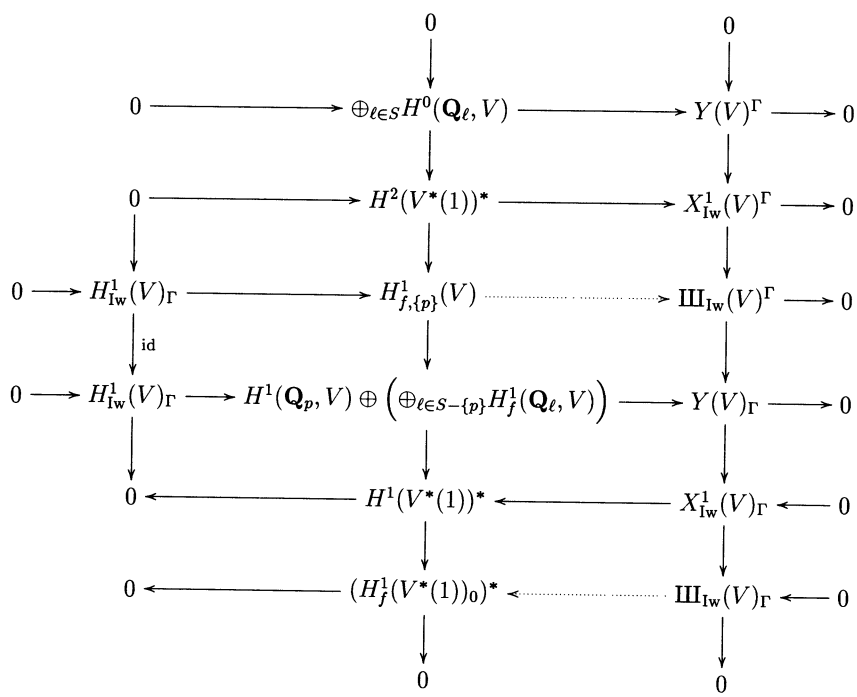
Finalement, pour passer de $\Lambda x_1 + \cdots + \Lambda x_{d^-}$ à $H^1_{\mathrm{Iw}}(T)$ et de $\Lambda z_1 + \cdots + \Lambda z_{d^-}$ à $H^1_{\mathrm{Iw}}(\mathbf{Q}_p, T)$, on utilise le fait que le sous-$\mathbf{Z}_p$-module de $H^1(T)_u = H^1_{\mathrm{Iw}}(T)_\Gamma$ engendré par $\int_\Gamma x_1, \ldots, \int_\Gamma x_{d^-}$ est d'indice $\frac{|H(x)_\Gamma|}{|H(x)^\Gamma|} = \int_\Gamma \mu_{H(x)}$ dans $H^1(T)_u$ et le sous-$\mathbf{Z}_p$-module de $H^1_{\mathrm{Iw}}(\mathbf{Q}_p, T)_\Gamma$ engendré par $\int_\Gamma z_1, \ldots, \int_\Gamma z_{d^-}$ est d'indice $\frac{|H(z)_\Gamma|}{|H(z)^\Gamma|} = \int_\Gamma \mu_{H(z)}$.

3.8. Étude de μ_T^2

Pour démontrer la proposition 3.5, on introduit le noyau $Ш_{\mathrm{Iw}}(T)$ de l'application de $H^2_{\mathrm{Iw}}(T)$ dans $\oplus_{\ell\in S} H^2_{\mathrm{Iw}}(\mathbf{Q}_\ell, T)$. On a alors $\mu_T^2 = \mu_{Ш_{\mathrm{Iw}}(T)}(\mu_{X^0_{\mathrm{Iw}}(T)})^{-1}$. D'autre part, d'après le lemme 3.7, la fonction $\int_\Gamma \langle\chi_{\mathrm{cycl}}\rangle^s \mu_{Ш_{\mathrm{Iw}}(T)}$ a un zéro d'ordre au moins égal à $r_{p,2}(V) = \dim Ш_{\mathrm{Iw}}(V)_\Gamma = \dim H^1_f(V^*(1))_0$ et on a

$$\lim_{s\to 0} s^{-r_{p,2}(V)} \int_\Gamma \langle\chi_{\mathrm{cycl}}\rangle^s \mu_{Ш_{\mathrm{Iw}}(T)} \sim_p \log\chi(\gamma)^{r_{p,2}(V)} [Ш_{\mathrm{Iw}}(T)_\Gamma : Ш_{\mathrm{Iw}}(T)^\Gamma].$$

Le problème est donc d'arriver à calculer $[Ш_{\mathrm{Iw}}(T)_\Gamma : Ш_{\mathrm{Iw}}(T)^\Gamma]$. Pour ce faire, considérons le diagramme suivant.



Dans ce diagramme, la colonne du milieu est obtenue en tronquant une suite exacte de Poitou-Tate modifiée et les lignes, dont les définitions sont données ci-dessous, sont exactes.

Les seconde et cinquième lignes proviennent du diagramme récapitulatif et elles sont exactes car on a supposé $X^2_{\mathrm{Iw}}(T)=0$ (conjecture de Leopoldt faible) et $\mathrm{D}_{\mathrm{cris}}(V)^{\varphi=p^{-1}}=0$, ce qui entraîne $H^0(\mathbf{Q}_p,V^*(1))=0$ et $X^0_{\mathrm{Iw}}(V)^\Gamma=0$.

Les première et quatrième lignes proviennent de la suite exacte $0\to H^1_{\mathrm{Iw}}(V)\to \oplus_{\ell\in S}H^1_{\mathrm{Iw}}(\mathbf{Q}_\ell,V)\to Y(V)\to 0$ via le lemme suivant [39] qui permet aussi de démontrer leur exactitude (si on utilise en plus le fait que $H^1_{\mathrm{Iw}}(V)^\Gamma=H^0(V)=0$ à cause de l'hypothèse $\mathrm{D}_{\mathrm{cris}}(V)^{\varphi=1}=0$).

LEMME 3.8. — i) Si $\ell\neq p$, alors $H^1_{\mathrm{Iw}}(\mathbf{Q}_\ell,T)$ est d'indice fini égal à $\mathrm{Tam}^0_\ell(T)$ dans $H^1_f(\mathbf{Q}_\ell,T)$.

ii) $H^1_{\mathrm{Iw}}(\mathbf{Q}_p,T)$ est d'indice fini dans $H^1(\mathbf{Q}_p,T)$.

iii) L'application naturelle de $H^1(T)_u$ dans $\oplus_{\ell\in S}H^1_{\mathrm{Iw}}(\mathbf{Q}_\ell,T)$ est injective.

Le diagramme formé des flèches continues étant commutatif et les colonnes étant exactes, il existe une unique flèche (pointillée) de $Ш_{\mathrm{Iw}}(V)_\Gamma$ dans $(H^1_f(V^*(1))_0)^*$ laissant le diagramme commutatif. Finalement, le lemme 3.8 montre que $H^1(V)_u$ est inclus dans $H^1_{f,\{p\}}(V)$ et, si on utilise le diagramme commutatif

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \oplus_{\ell\in S}H^1_{\mathrm{Iw}}(\mathbf{Q}_\ell,V)_\Gamma & \longrightarrow & \oplus_{\ell\in S}H^1(\mathbf{Q}_\ell,V) & \longrightarrow & \oplus_{\ell\in S}H^2_{\mathrm{Iw}}(\mathbf{Q}_\ell,V)^\Gamma & \longrightarrow & 0\\
 & & \uparrow & & \uparrow & & \uparrow & & \\
0 & \longrightarrow & H^1_{\mathrm{Iw}}(V)_\Gamma & \longrightarrow & H^1(V) & \longrightarrow & H^2_{\mathrm{Iw}}(V)^\Gamma & \longrightarrow & 0,
\end{array}$$

on montre que $H^1_{f,\{p\}}(V)\subset H^1(V)$ est l'image réciproque de $Ш_{\mathrm{Iw}}(V)^\Gamma\subset H^2_{\mathrm{Iw}}(V)^\Gamma$ par l'application naturelle de $H^1(V)$ dans $H^2_{\mathrm{Iw}}(V)^\Gamma$, ce qui permet de définir la flèche en pointillé de $H^1_{f,\{p\}}(V)$ dans $Ш_{\mathrm{Iw}}(V)^\Gamma$.

PROPOSITION 3.9. — *Le diagramme ainsi obtenu est commutatif et l'application de $H^1_f(V)$ dans $(H^1_f(V^*(1))_0)^*$ obtenue en composant les applications*

$$H^1_f(V)\to Ш_{\mathrm{Iw}}(V)^\Gamma\to Ш_{\mathrm{Iw}}(V)_\Gamma\to (H^1_f(V^*(1))_0)^*,$$

coïncide avec celle induite par l'accouplement "hauteur p-adique" divisée par $\log\chi(\gamma)$.

Cette proposition se démontre [39] en revenant à la définition de la suite exacte de Poitou-Tate. On a d'autre part le diagramme ci-dessous qui est la version entière du diagramme permettant de définir les flèches $H^1_f(V)\to Ш_{\mathrm{Iw}}(V)^\Gamma$ et $Ш_{\mathrm{Iw}}(V)_\Gamma\to(H^1_f(V^*(1))_0)^*$ et dans lequel $Z=\frac{H^1(\mathbf{Q}_p,T)}{H^1_{\mathrm{Iw}}(\mathbf{Q}_p,T)_\Gamma}\oplus\left(\oplus_{\ell\in S-\{p\}}\frac{H^1_f(\mathbf{Q}_\ell,T)}{H^1_{\mathrm{Iw}}(\mathbf{Q}_\ell,T)_\Gamma}\right)$ est un groupe fini ayant pour cardinal $|Z|=C_p\prod_{\ell\in S-\{p\}}\mathrm{Tam}^0_\ell(T)$ d'après le lemme 3.8. On déduit de la proposition 3.9 et du diagramme que $[Ш_{\mathrm{Iw}}(T)_\Gamma:Ш_{\mathrm{Iw}}(T)^\Gamma]$ est aussi égal à

$$\log(\chi(\gamma))^{-r_{p,2}(V)}\left\langle(H^1_f(T^\vee(1))_0)^\vee,\frac{H^1_f(T)}{H^1_f(T)_u}\right\rangle_{\mathrm{ht}}\left[\frac{H^1_{f,\{p\}}(T)}{H^1(T)_u}:\frac{H^1_f(T)}{H^1_f(T)_u}\right]^{-1}\frac{|Z|}{|X^0_{\mathrm{Iw}}(T)^\Gamma|},$$

ce qui, compte tenu des formules $\int_\Gamma \mu_{X^0_{\mathrm{Iw}}(T)} = \frac{|X^0_{\mathrm{Iw}}(T)_\Gamma|}{|X^0_{\mathrm{Iw}}(T)^\Gamma|}$ et $X^0_{\mathrm{Iw}}(T)_\Gamma = H^0(T^\vee(1))^\vee$ permet de terminer la démonstration de la proposition 3.5.

$$
\begin{array}{ccccccccc}
 & & & & 0 & & 0 & & \\
 & & & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & \oplus_{\ell\in S} H^0(\mathbf{Q}_\ell, T) & \longrightarrow & & & Y(T)^\Gamma & \longrightarrow & 0 \\
 & & \downarrow & & & & \downarrow & & \\
0 & \longrightarrow & H^2(T^\vee(1))^\vee & \longrightarrow & & & X^1_{\mathrm{Iw}}(T)^\Gamma & \longrightarrow & 0 \\
 & & \downarrow & & & & \downarrow & & \\
 & & H^1_{f,\{p\}}(T)/H^1(T)_u & \dashrightarrow & & & Ш_{\mathrm{Iw}}(T)^\Gamma & & \\
 & & \downarrow & & & & \downarrow & & \\
0 & \longleftarrow & Z & \longleftarrow & \frac{H^1(\mathbf{Q}_p,T)\oplus\left(\oplus_{\ell\in S-\{p\}} H^1_f(\mathbf{Q}_\ell,T)\right)}{H^1_{\mathrm{Iw}}(T)_\Gamma} & \longleftarrow & Y(T)_\Gamma & \longleftarrow & 0 \\
 & & \downarrow & & & & \downarrow & & \\
0 & \longleftarrow & X^0_{\mathrm{Iw}}(T)^\Gamma & \longleftarrow & H^1(T^\vee(1))^\vee & \longleftarrow & X^1_{\mathrm{Iw}}(T)_\Gamma & \longleftarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & & & (H^1_f(T^\vee(1))_0)^\vee & \dashleftarrow & Ш_{\mathrm{Iw}}(T)_\Gamma & & \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & 0 & & 0 & &
\end{array}
$$

RÉFÉRENCES

[1] Y. Amice, Interpolation p-adique, Bull. Soc. France **92**, 117-180, 1964

[2] Y. Amice, Duals. Proc. of a conf. on p-adic analysis (Nijmegen 1978), 1-15, Nijmegen, Math. Institut Katholische Univ., 1978

[3] Y. Amice et J. Vélu, Distributions p-adiques associées aux séries de Hecke, Astérisque **24-25**, Soc. Math. France, Paris 119–131, 1975

[4] D. Barsky, Fonctions k-lipschitziennes sur un anneau local et polynômes à valeurs entières, Bull. Soc. Math. France **101**, 397-411, 1973

[5] D. Benois, On Iwasawa theory of cristalline representations, preprint 1998

[6] A. Beilinson, Higher regulators and values of L-functions, J. Soviet Math. **30**, 2036-2070, 1985

[7] A. Beilinson, Polylogarithm and cyclotomic elements, preprint, 1989

[8] A. Beilinson et P. Deligne, Motivic polylogarithm and Zagier conjecture, preprint 1992

[9] S. Bloch et K. Kato, L functions and Tamagawa numbers of motives, dans The Grothendieck Festschrift, vol. 1, Prog. in Math. **86**, Birkhaüser, Boston, 333–400, 1990

[10] F. Cherbonnier, P. Colmez, Théorie d'Iwasawa des représentations p-adiques d'un corps local, J. Amer. Math. Soc **12**, 241-268, 1999

[11] J. Coates, p-adic L functions, Séminaire Bourbaki (1988), exposé 701, Astérisque **177-178**, 33-59, 1989

[12] J. Coates et A. Wiles, On p-adic L-functions and elliptic units, J. Australian Math. Soc., **A 26**, 1-25, 1978

[13] R. Coleman, Division values in local fields, Inv. Math. **53**, 91-116, 1979

[14] P. Colmez, Théorie d'Iwasawa des représentations de de Rham d'un corps local, Ann. Math. **148**, 485-571, 1998

[15] P. Deligne, Valeurs de fonctions L et périodes d'intégrales, dans Automorphic forms, representations and L-functions, Proc. Symp. Pure Math., American Math. Soc., **33**, 313-346, 1979

[16] J.-M. Fontaine, Sur certains types de représentations p-adiques du groupe de Galois d'un corps local ; construction d'un anneau de Barsotti-Tate, Ann. Math. **115**, 529-577, 1982

[17] J.-M. Fontaine, Le corps des périodes p-adiques. dans "*Périodes p-adiques*" exposé II, Astérisque **223**, 59-102, 1994

[18] J.-M. Fontaine, Représentations p-adiques des corps locaux, dans "*The Grothendieck Festschrift*", vol II, Birkhauser, Boston 249-309, 1991.

[19] J.-M. Fontaine, Valeurs spéciales de fonctions L des motifs, Séminaire Bourbaki, exposé **751**, Astérisque **206**, 205-249, 1992

[20] J.-M. Fontaine et B. Mazur, Geometric Galois representations, dans *Elliptic curves, modular forms and Fermat's last theorem (Hong-Kong, 1993)*, Ser. Number Theory **I**, Internat. Press, Cambridge MA, 41-78, 1995

[21] J.-M. Fontaine et B. Perrin-Riou, Autour des conjectures de Bloch et Kato : cohomologie galoisienne et valeurs de fonctions L, dans *Motives (Seattle)*, Proceedings of Symposia in Pure Mathematics, vol. **55**, part 1, 599-706, 1994

[22] R. Greenberg, Iwasawa theory for p-adic representations, Adv. Stud. Pure Math. **17**, 97-137, 1989

[23] H. Hida, *Elementary theory of L-functions and Eisenstein series*, London Math. Soc. Stud. Texts **26**, Cambridge University Press, 1993

[24] H. Hida, On the search of genuine p-adic modular L-functions for GL(n), Mém. Soc. Math. Fr. **67**, 1996

[25] A. Huber et G. Kings, Degeneration of ℓ-adic Eisenstein classes and of the elliptic polylog. Invent. Math. **135**, 545–594, 1999

[26] A. Huber et J. Wildeshaus, Classical motivic polylogarithm according to Beilinson and Deligne, Doc. Math. **3**, 27-133 et 297-299, 1998

[27] U. Jannsen, On the l-adic cohomology of varieties over number fields and its Galois cohomology, *Galois groups over* **Q**, Math. Sci. Res. Inst. Publ. **16**, 315-360, Springer, 1989

[28] K. Kato, Lectures on the approach to Iwasawa theory for Hasse-Weil L-functions via B_{dR}, dans Arithmetic Algebraic Geometry, Lecture Notes in Math. **1553**, 1993.

[29] K. Kato, Euler systems, Iwasawa theory and Selmer groups, preprint 1995

[30] K. Kato, M. Kurihara and T. Tsuji, Local Iwasawa theory of Perrin-Riou and syntomic complexes, preprint 1996

[31] V. Kolyvagin, Euler systems, The Grothendieck Festschrift, vol. **2**, Prog. in Math. **87**, 436-483, Birkhaüser, Boston, 1990

[32] Y. Manin, Periods of cusp forms, and p-adic Hecke series, Math USSR-Sb. **92**, 371-393, 1973

[33] B. Mazur, J. Tate et J. Teitelbaum, On p-adic analogues of the conjectures of Birch and Swinnerton-Dyer, Invent. Math. **84**, 1-48, 1986

[34] B. Mazur et A. Wiles, Class fields of abelian extensions of **Q**, Invent. Math. **76**, 179-330, 1984

[35] J. Milne, *Arithmetic duality theorems,* Perspectives in Mathematics, vol. **1**, Academic Press, 1986.

[36] J. Nekovář, On p-adic height pairings, dans Séminaire de théorie des nombres 1990-1991, édité par S. David, 127-202, Birkhaüser Boston, 1993

[37] J. Nekovář, Beilinson's conjectures, dans *Motives (Seattle)*, Proceedings of Symposia in Pure Mathematics, vol. **55**, part 1, 537-570, 1994

[38] B. Perrin-Riou, Théorie d'Iwasawa p-adique locale et globale, Invent. Math. **99**, 247-292, 1990

[39] B. Perrin-Riou, Théorie d'Iwasawa et hauteurs p-adiques, Invent. Math. **109**, 137-185, 1992

[40] B. Perrin-Riou, Théorie d'Iwasawa des représentations p-adiques sur un corps local, Invent. Math. **115** 81-149, 1994

[41] B. Perrin-Riou, *Fonctions L p-adiques des représentations p-adiques*, Astérisque **229**, 1995

[42] B. Perrin-Riou, Fonctions L p-adiques, dans *Proc. Int. Congress Math. Zürich*, 400-410, Birkhäuser, 1995

[43] B. Perrin-Riou, Systèmes d'Euler et représentations p-adiques, Ann. Inst. Fourier **48**, 1231-1307, 1998

[44] K. Rubin, The main conjecture, Appendix to Cyclotomic fields (seconde édition) par S. Lang, Graduate Texts in Math. **121**, Springer-Verlag, 1990.

[45] K. Rubin, Euler systems, Hermann Weyl lecture, livre en préparation

[46] A. Scholl, Remarks on special values of L-functions, dans *L-functions and Arithmetic*, Proc. of the Durham Symp., London Math. Soc. L.N.S. 153, Cambridge University Press, 1991, 373-392

[47] A. Scholl, Height pairings and special values of L-functions, dans *Motives (Seattle)*, Proceedings of Symposia in Pure Mathematics, vol. **55**, part 1, 571-598, 1994

[48] A. Scholl, An introduction to Kato's Euler systems, dans *Galois representations in arithmetic algebraic geometry*, 379-460, Cambridge University Press, 1998

[49] P. Schneider, p-adic height pairings, II, Invent. Math. **79**, 329-374, 1985

[50] J.-P. Serre, Sur le résidu de la fonction zêta p-adique d'un corps de nombres, C. R. Acad. Sci. Paris **287** série A, 83-126, 1978

[51] C. Soulé, Eléments cyclotomiques en K-théorie, Astérisque **147-148**, 225-257, 1987

[52] J. Tate, p-divisible groups, Proc. of a conference on local fields, Nuffic Summer School at Driebergen, Springer, Berlin, 158-183, 1967

[53] M. Vishik, Non-archimedian measures connected with Dirichlet series, Math. USSR Sbornik **28**, 216-228, 1976

Pierre COLMEZ

École Normale supérieure
45 rue d'Ulm
75005 PARIS

Institut de mathématiques de Jussieu
4 place Jussieu
75005 PARIS
E-mail : `colmez@ens.fr`

Astérisque

JEAN-PIERRE DEMAILLY

Méthodes L^2 et résultats effectifs en géométrie algébrique

Astérisque, tome 266 (2000), Séminaire Bourbaki, exp. n° 852, p. 59-90
<http://www.numdam.org/item?id=SB_1998-1999__41__59_0>

© Société mathématique de France, 2000, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (http://smf4.emath.fr/Publications/Asterisque/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM
Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Séminaire BOURBAKI Novembre 1998
51ème année, 1998-99, n° 852

MÉTHODES L^2 ET RÉSULTATS EFFECTIFS EN GÉOMÉTRIE ALGÉBRIQUE

par **Jean-Pierre DEMAILLY**

1. INTRODUCTION

La théorie des systèmes linéaires adjoints a pour but d'étudier les espaces de sections $H^0(X, K_X + mL)$ associés à un fibré en droites L ample – ou du moins suffisamment positif – sur une variété algébrique projective X de dimension n. La motivation principale est la construction de plongements d'une "variété algébrique polarisée" (X, L) donnée dans un espace projectif complexe $\mathbb{P}^N_{\mathbb{C}}$, avec des bornes effectives explicites pour les degrés. À leur tour, de tels plongements peuvent être utilisés pour démontrer des théorèmes de finitude ou pour essayer de classifier les structures algébriques sur une variété de type topologique donné.

On supposera tout au long de cet exposé que X est lisse, définie sur $\mathbb{C}$, et on notera $K_X = \Lambda^n T^\star_X$, $n = \dim X$, le *fibré en droites canonique* de X. On utilisera la notation additive usuelle pour le groupe de Picard: $K_X + mL$ est donc synonyme de $K_X \otimes L^{\otimes m}$. L'une des questions les plus motivantes pour la théorie des systèmes linéaires adjoints a sans doute été la conjecture suivante, due à T. Fujita [Fuj87, 88].

Conjecture (Fujita). *Si L est un fibré en droites ample sur une variété projective X de dimension n, alors*

(i) *$K_X + mL$ est engendré par ses sections pour $m \geq n + 1$.*

(ii) *$K_X + mL$ est très ample pour $m \geq n + 2$.*

À ce jour, la partie (ii) de la conjecture de Fujita semble encore hors de portée hormis le cas bien compris des dimensions 1 et 2 (cf. I. Reider [Rei87]), mais la partie (i) a fait l'objet de nombreux travaux qui ont conduit à une réponse positive jusqu'en dimension 5 (Ein-Lazarsfeld [EL93] en dimension 3, Y. Kawamata [Kaw97a] en dimension 4, S. Helmke [Hel98] en dimension 5).

Il faut observer que ce type de résultat, fournissant une borne universelle ne dépendant que de la dimension n pour l'entier m, se saurait être vrai pour les systèmes linéaires $H^0(X, mL)$. En effet, si X est une courbe de genre g, il est bien connu que $H^0(X, mL)$ n'a pas de sections si $m < g$, pour L générique de degré 1. Par ailleurs, les bornes de la conjecture de Fujita sont déjà optimales dans le cas où $X = \mathbb{P}^n_{\mathbb{C}}$, $L = \mathcal{O}(1)$, puisqu'on a alors $K_X = \mathcal{O}(-n-1)$.

Le caractère inévitable du fibré canonique s'explique par son intervention dans les théorèmes d'annulation fondamentaux tels que le théorème de Kodaira-Nakano, le théorème de Kawamata-Viehweg ou le théorème de Nadel (voir § 2). L'approche présentée ici s'appuira sur une étude approfondie des métriques singulières et des idéaux multiplicateurs de Nadel, qui mesurent de façon précise l'influence des ensembles-base des sytèmes linéaires considérés. Le cadre de travail est la théorie de Hodge L^2, et les outils analytiques sous-jacents sont les théorèmes d'existence L^2 pour les solutions de l'opérateur $\overline{\partial}$. Une des applications marquantes de ces techniques est la démonstration du théorème de l'invariance des plurigenres par déformation, récemment obtenu par Y.T. Siu [Siu97] dans le cas des variétés de type général. Le lecteur pourra consulter Y. Kawamata [Kaw97b, Kaw98] pour diverses généralisations dans un contexte plus algébrique, incluant notamment le cas des déformations de variétés ayant des singularités canoniques.

Théorème (Siu). *Soit $\mathcal{X} \to S$ une famille projective lisse de variétés de type général au-dessus d'une base S irréductible. Alors les plurigenres $p_m(X_t) = h^0(X_t, mK_{X_t})$ des fibres sont indépendants de t pour tout $m \geq 0$.*

L'assertion plus générale où les fibres X_t seraient de dimension de Kodaira quelconque est encore conjecturale (et nécessite vraisemblablement des techniques de théorie de Hodge beaucoup plus élaborées). Nous allons maintenant donner un aperçu des méthodes utilisées et de la preuve des principaux résultats – en essayant de nous adresser au lecteur non nécessairement spécialiste de la géométrie algébrique.

2. MÉTRIQUES SINGULIÈRES ET THÉORÈMES D'ANNULATION

2.1. Métriques hermitiennes singulières

Soit (L, h) un fibré holomorphe en droites hermitien sur une variété complexe X. On ne suppose pas a priori que la métrique h soit de classe C^∞, mais on pose toutefois une condition restrictive de manière à pouvoir calculer la courbure au sens des courants (cf. [Dem90, DPS94]).

2.1.1. Définition. *Une métrique (hermitienne) singulière sur un fibré en droites L est une métrique donnée dans toute trivialisation $\tau : L_{|U} \xrightarrow{\simeq} U \times \mathbb{C}$ par*

$$\|\xi\| = |\tau(\xi)|\, e^{-\varphi(x)}, \qquad x \in U,\ \xi \in L_x$$

où $\varphi \in L^1_{\mathrm{loc}}(U)$ est une fonction arbitraire localement intégrable (pour la mesure de Lebesgue dans des coordonnées locales), appelée poids de la métrique par rapport à la trivialisation τ.

Si $\tau' : L_{|U'} \to U' \times \mathbb{C}$ est une autre trivialisation, φ' le poids associé sur U' et $g \in \mathcal{O}^*(U \cap U')$ la fonction de transition, alors $\tau'(\xi) = g(x)\,\tau(\xi)$ pour tout $\xi \in L_x$, donc $\varphi' = \varphi + \log|g|$ sur $U \cap U'$. Une définition possible de la *forme de courbure* de L consiste à poser

$$\Theta_h(L) = \frac{i}{\pi}\partial\overline{\partial}\varphi \tag{2.1.2}$$

sur U. C'est une 2-forme réelle d-fermée de type $(1,1)$. La formule $\varphi' = \varphi + \log|g|$ montre précisément que $i\partial\overline{\partial}\varphi$ est invariant par changement de trivialisation, et par conséquent $\Theta_h(L)$ est un *courant* de type $(1,1)$ globalement défini sur X (rappelons que, d'après G. De Rham [Rh55], un courant est simplement une forme différentielle à coefficients distributions); l'hypothèse $\varphi \in L^1_{\mathrm{loc}}(U)$ garantit en effet que $\Theta(L)$ existe au sens des distributions. Un changement de métrique $h \mapsto h'$ s'obtient par $h' = h\,e^{-\psi}$ avec $\psi \in L^1_{\mathrm{loc}}(X)$, de sorte que

$$\Theta_{h'}(L) = \Theta_h(L) + \frac{i}{\pi}\partial\overline{\partial}\psi \tag{2.1.3}$$

appartient à la même classe de cohomologie de De Rham que $\Theta_h(L)$ dans $H^2_{DR}(X,\mathbb{R})$. De plus, on sait (cf. par exemple [GH78]) que la première classe de Chern $c_1(L)$ est définie en cohomologie de De Rham précisément par le courant $\Theta_h(L)$. Rappelons qu'un courant réel $T = i\sum_{1\le j,k\le n} T_{jk}\,dz_j \wedge d\overline{z}_k$ de type $(1,1)$ est dit *(semi-)positif* si $\sum_{1\le j,k\le n}\lambda_j\overline{\lambda}_k T_{jk}$ est une mesure positive pour tout système de coefficients complexes $\lambda = (\lambda_j) \in \mathbb{C}^n$. Une fonction $\varphi \in L^1_{\mathrm{loc}}$ est dite *plurisousharmonique* si $i\partial\overline{\partial}\varphi = i\sum \partial^2\varphi/\partial z_j\partial\overline{z}_k\, dz_j \wedge d\overline{z}_k \ge 0$. On introduit la définition suivante.

2.1.4. Définition. *Le fibré hermitien singulier (L,h) est dit à courbure semi-positive (resp. définie positive) si le $(1,1)$-courant de courbure $\Theta_h(L)$ est semi-positif, resp. si $\Theta_h(L)$ est défini positif, i.e. il existe une $(1,1)$-forme $\omega = i\sum_{1\le j,k\le n}\omega_{jk}(z)dz_j \wedge d\overline{z}_k$ de classe C^∞, définie positive, et $\varepsilon > 0$ tels que $\Theta_h(L) \ge \varepsilon\omega$.*

Avant d'aller plus loin, nous discutons deux exemples fondamentaux.

2.1.5. Exemple. Soit $D = \sum \alpha_j D_j$ un diviseur à coefficients $\alpha_j \in \mathbb{Z}$ et soit $L = \mathcal{O}(D)$ le faisceau inversible associé, défini comme le faisceau des fonctions méromorphes u telles que $\mathrm{div}(u) + D \geq 0$; le fibré en droites correspondant peut alors être muni de la métrique singulière définie par $\|u\| = |u|$ (module de la fonction méromorphe u). Si g_j est un générateur de l'idéal de D_j sur un ouvert $U \subset X$, alors $\tau(u) = u \prod g_j^{\alpha_j}$ définit une trivialisation de $\mathcal{O}(D)$ sur U, donc notre métrique singulière est associée au poids $\varphi = \sum \alpha_j \log |g_j|$. L'équation de Lelong-Poincaré ([Lel57, 69]) implique

$$\Theta\big(\mathcal{O}(D)\big) = \frac{i}{\pi}\partial\overline{\partial}\varphi = [D],$$

où $[D] = \sum \alpha_j [D_j]$ désigne le courant d'intégration sur D. La courbure est semi-positive au sens des courants si et seulement si le diviseur D est effectif (i.e. à coefficients $\alpha_j \geq 0$).

2.1.6. Exemple. Supposons que $\sigma_1, \ldots, \sigma_N$ soient des sections holomorphes non nulles de L. On peut alors définir une métrique hermitienne naturelle h (éventuellement singulière) sur $L^\star$, en posant

$$\|\xi^\star\|^2 = \sum_{1 \leq j \leq n} \big|\xi^\star . \sigma_j(x)\big|^2 \quad \text{pour} \quad \xi^\star \in E_x^\star.$$

La métrique duale de L est donnée par

$$\|\xi\|^2 = \frac{|\tau(\xi)|^2}{|\tau(\sigma_1(x))|^2 + \cdots + |\tau(\sigma_N(x))|^2}$$

par rapport à toute trivialisation locale τ. La fonction poids associée est donc donnée par $\varphi(x) = \log\big(\sum_{1 \leq j \leq N} |\tau(\sigma_j(x))|^2\big)^{1/2}$. Notons

$$\Sigma = |\sigma_1, \ldots, \sigma_N| := \mathrm{Vect}(\sigma_1, \ldots, \sigma_N)$$

le système linéaire défini par $\sigma_1, \ldots, \sigma_N$ et $B_\Sigma = \bigcap \sigma_j^{-1}(0)$ son ensemble base. On a une application méromorphe

$$\Phi_\Sigma : X \smallsetminus B_\Sigma \to P(\Sigma^\star) \simeq \mathbb{P}^{N-1}, \qquad x \mapsto \Phi_\Sigma(x) \simeq [\sigma_1(x) : \sigma_2(x) : \cdots : \sigma_N(x)],$$

qui à x associe l'hyperplan

$$\Phi_\Sigma(x) = \big\{\sigma = \sum \xi_j \sigma_j \,;\ \sigma(x) = \sum \xi_j \sigma_j(x) = 0\big\} \subset \Sigma.$$

Avec ces notations, la courbure $\Theta_h(L)$ restreinte à $X \setminus B_\Sigma$ s'identifie à l'image-inverse par Φ_Σ de la métrique de Fubini-Study $\omega_{\mathrm{FS}} = \frac{i}{2\pi}\partial\overline{\partial}\log(|z_1|^2 + \cdots + |z_N|^2)$ sur $\mathbb{P}^{N-1}$. Le courant $\Theta_h(L)$, qui est de bidimension $(n-1, n-1)$, ne peut porter de masse sur B_Σ lorsque B_Σ est de codimension ≥ 2, mais il peut se produire que B_Σ ait une composante divisorielle égale au pgcd D des diviseurs $\sigma_j = 0$. Dans ce cas, on vérifie aisément que $\Theta_h(L)$ est égal au courant d'intégration $[D]$ en restriction à B_Σ. Dans tous les cas, $\Theta_h(L)$ est un courant positif fermé.

2.2. Identité de Bochner-Kodaira-Nakano

Dans cette section, on désigne par (L, h) un fibré holomorphe hermitien au-dessus d'une variété complexe X, tel que la métrique h soit de classe C^∞. Rappelons tout d'abord le lemme classique suivant (voir par exemple [GH78]).

2.2.1. Lemme. *Il existe une unique connexion $D = \nabla + \overline{\nabla}$ sur L, dite connexion de Chern, ayant les propriétés suivantes:*

(i) *D opère sur les sections C^∞ des fibrés $\Lambda^{\bullet,\bullet}T^\star_X \otimes L$, la composante ∇ (resp. $\overline{\nabla}$) envoyant les formes de type (p,q) dans les formes de type $(p+1,q)$, (respectivement $(p,q+1)$).*

(ii) *D satisfait la règle de Leibnitz, à savoir $D(f \wedge u) = df \wedge \nabla u + (-1)^{\deg f} f \wedge \nabla u$, si f est une forme à valeurs scalaires et u une forme à valeurs dans L.*

(iii) *D est "holomorphe", i.e. $\overline{\nabla} = \overline{\partial}$.*

(iv) *D est hermitienne, i.e. la métrique h est une section parallèle du fibré des matrices hermitiennes* $\mathrm{Herm}(E) \subset E^\star \otimes \overline{E}^\star$.

Si $L_{|U} \simeq U \times \mathbb{C}$ est localement trivialisé et si la métrique h donnée par un poids $e^{-\varphi}$, on vérifie facilement que

$$\nabla u = \partial u - 2\partial\varphi \wedge u, \qquad \overline{\nabla} u = \overline{\partial} u, \qquad D^2 u = (\nabla\overline{\nabla} + \overline{\nabla}\nabla)u = 2\partial\overline{\partial}\varphi \wedge u, \tag{2.2.2}$$

en sorte que $\frac{i}{2\pi}D^2 u = \Theta_h(L) \wedge u$. On suppose maintenant que X est munie d'une métrique hermitienne $\omega = i\sum_{1\leq j,k\leq n} \omega_{jk}(z)\, dz_j \wedge d\overline{z}_k$ définie positive. Une telle métrique permet de définir des normes L^2 et des espaces de sections L^2 globales $L^2(X, \Lambda^{p,q}T^\star_X \otimes L)$ en posant

$$\|u\|^2 = \|u\|^2_{\omega,h} = \int_X |u|^2_\omega e^{-2\varphi} dV_\omega,$$

où $dV_\omega = \frac{1}{n!}w^n$ est l'élément de volume kählérien, $|u|_\omega$ la norme ponctuelle induite par ω sur $\Lambda^{p,q}T^\star_X$ et $e^{-\varphi}$ le poids de la métrique h sur L (il y a quelque abus dans

cette notation, car le poids φ n'est pas global, mais on utilisera tout de même cette notation par souci de simplicité). La norme L^2 permet de définir des adjoints formels $\nabla^\star$ et $\overline{\nabla}^\star$, de types respectifs $(-1,0)$ et $(0,-1)$, et on considère les opérateurs de Laplace-Beltrami associés

$$\Delta = DD^\star + D^\star D, \qquad \Box = \nabla\nabla^\star + \nabla^\star\nabla, \qquad \overline{\Box} = \overline{\nabla}\,\overline{\nabla}^\star + \overline{\nabla}^\star\overline{\nabla}.$$

On a alors l'identité fondamentale suivante.

2.2.3. Identité de Bochner-Kodaira-Nakano. *Si ω est kählérienne, les laplaciens Δ, $\Box$, $\overline{\Box}$ vérifient $\Delta = \Box + \overline{\Box}$ et*

$$\overline{\Box} = \Box + [\Theta_h(L), \Lambda]$$

où Λ est l'adjoint de l'opérateur $s \mapsto \omega \wedge s$, $\Theta_h(L)$ l'opérateur de multiplication par le tenseur de courbure de (L,h) et $[\bullet,\bullet]$ le crochet de commutation.

En chaque point $x \in X$, on peut choisir un système de coordonnées $(z_1, \ldots, z_n)$ qui diagonalise simultanément les formes hermitiennes $\omega(x)$ et $\Theta_h(L)(x)$, de telle manière que

$$\omega(x) = i\sum_{1\le j\le n} dz_j \wedge d\overline{z}_j, \qquad \Theta_h(L)(x) = i\sum_{1\le j\le n} \gamma_j dz_j \wedge d\overline{z}_j$$

avec $\gamma_1 \le \cdots \le \gamma_n$. Les valeurs propres de courbure $\gamma_j = \gamma_j(x)$ sont alors définies de manière unique et continues en x. Pour toute (p,q)-forme $u = \sum u_{JK} dz_J \wedge d\overline{z}_K \otimes s$ (avec s section holomorphe trivialisante de L, $|s|_h = e^{-\varphi}$), un calcul explicite donne

$$\begin{aligned}\langle [i\,\Theta_h(L), \Lambda]u, u\rangle &= \sum_{|J|=p,\,|K|=q} \Big(\sum_{j\in J}\gamma_j + \sum_{j\in K}\gamma_j - \sum_{1\le j\le n}\gamma_j\Big)|u_{JK}|^2 e^{-2\varphi} \\ &\ge (\gamma_1 + \cdots + \gamma_q - \gamma_{n-p+1} - \cdots - \gamma_n)|u|^2 e^{-2\varphi}. \qquad (2.2.4)\end{aligned}$$

Comme $\langle\!\langle\, \overline{\Box}u, u\,\rangle\!\rangle = \|\overline{\nabla}u\|^2 + \|\overline{\nabla}^\star u\|^2$, on déduit de (2.2.3) et (2.2.4) l'inégalité fondamentale

$$\|\overline{\nabla}u\|^2 + \|\overline{\nabla}^\star u\|^2 \ge \int_X (\gamma_1 + \cdots + \gamma_q - \gamma_{n-p+1} - \cdots - \gamma_n)|u|^2 e^{-\varphi} dV_\omega. \qquad (2.2.5)$$

Supposons que $\Theta_h(L)$ soit positive. Dans ce cas, il est naturel de munir X de la métrique kählérienne particulière $\omega = \Theta_h(L)$. Alors $\gamma_j = 1$ pour $j = 1, 2, \ldots, n$ et on obtient l'égalité $\|\overline{\nabla}u\|^2 + \|\overline{\nabla}^\star u\|^2 = (p+q-n)\|u\|^2$. Ceci implique en particulier

que les formes $\overline{\Box}$-harmoniques sont nulles en tout bidegré (p,q) tel que $p+q>n$, d'où le

2.2.6. Théorème d'annulation de Kodaira-Akizuki-Nakano. *Si (L,h) est un fibré en droites à courbure positive sur une variété complexe compacte X, alors*

$$H^{p,q}(X,L)=H^q(X,\Omega_X^p\otimes L)=0 \qquad \textit{pour}\ \ p+q>n=\dim X.$$

Observons cependant que le théorème de Kodaira-Akizuki-Nakano n'est pas valable dans le cas de fibrés munis de métriques h singulières (et en particulier, dans le cas de systèmes linéaires ayant un ensemble-base non trivial), car ω doit être lisse et le choix $\omega=\Theta_h(L)$ n'est pas permis dans ce cas.

2.3. Estimations L^2 de Hörmander-Andreotti-Vesentini

Dans cette section, on s'affranchit complètement de toute hypothèse de régularité sur la métrique h de L. On suppose simplement que $\Theta_h(L)$ est un courant positif. Les coefficients de ce courant sont alors des mesures, ses valeurs propres par rapport à la métrique de référence ω sont évaluées en considérant la partie absolument continue de $\Theta_h(L)$ par rapport à la mesure de Lebesgue (et en négligeant la partie singulière).

2.3.1. Théorème. *Soit (X,ω) une variété kählérienne, $\dim X=n$. On suppose que X est ou bien compacte ou bien faiblement pseudoconvexe, au sens où X possède une fonction d'exhaustion C^∞ faiblement plurisousharmonique. Soit (L,h) un fibré en droites holomorphe muni d'une métrique hermitienne singulière à courbure $\Theta_h(L)\geq 0$, et soient*

$$\gamma_1(x)\leq\cdots\leq\gamma_n(x)$$

les valeurs propres de courbure de $\Theta_h(L)$ par rapport à la métrique ω en tout point. Alors pour toute forme $g\in L^2(X,\Lambda^{n,q}T_X^\star\otimes E)$ telle que

$$\overline{\partial}g=0 \quad et \qquad \int_X(\gamma_1+\cdots+\gamma_q)^{-1}|g|^2e^{-2\varphi}\,dV_\omega<+\infty,$$

(on suppose donc $g(x)=0$ presque partout aux points où $\gamma_1(x)+\cdots+\gamma_q(x)=0$), il existe $f\in L^2(X,\Lambda^{n,q-1}T_X^\star\otimes E)$ telle que

$$\overline{\partial}f=g \quad et \qquad \int_X|f|^2e^{-2\varphi}\,dV_\omega\leq\int_X(\gamma_1+\cdots+\gamma_q)^{-1}|g|^2e^{-2\varphi}\,dV_\omega.$$

Preuve abrégée. Supposons d'abord que X soit compacte et que h soit de classe C^∞, en sorte que les valeurs propres γ_j sont des fonctions continues bornées. L'inégalité

(2.2.5) montre que pour $p = n$ on a

$$\|\overline{\nabla} u\|^2 + \|\overline{\nabla}^\star u\|^2 \geq \int_X (\gamma_1 + \cdots + \gamma_q)|u|^2 e^{-2\varphi} dV_\omega.$$

Les opérateurs $\overline{\nabla}$ et $\overline{\nabla}^\star$ peuvent être étendus en des opérateurs fermés à domaine dense sur les espaces de sections L^2, et l'inégalité ci-dessus s'étend à toute forme $u \in L^2$ dans l'intersection $\mathrm{Dom}(\overline{\nabla}) \cap \mathrm{Dom}(\overline{\nabla}^\star)$ de leurs domaines. Considérons la somme directe orthogonale

$$L^2(X, \Lambda^{n,q} T^\star_X \otimes L) = \mathrm{Ker}\, \overline{\nabla} \oplus (\mathrm{Ker}\, \overline{\nabla})^\perp, \qquad (\mathrm{Ker}\, \overline{\nabla})^\perp = \overline{\mathrm{Im}\, \overline{\nabla}^\star} \subset \mathrm{Ker}\, \overline{\nabla}^\star,$$

et écrivons un élément $u \in \mathrm{Dom}(\overline{\nabla}^\star)$ sous la forme $u = u_1 + u_2$ dans cette décomposition. Comme $g \in \mathrm{Ker}\, \overline{\nabla}$, on a $\langle\!\langle u, g \rangle\!\rangle = \langle\!\langle u_1, g \rangle\!\rangle$, d'où par Cauchy-Schwarz

$$\begin{aligned} \big|\langle\!\langle u, g \rangle\!\rangle\big|^2 &\leq \int_X (\gamma_1 + \cdots + \gamma_q)^{-1} |g|^2 e^{-2\varphi}\, dV_\omega \int_X (\gamma_1 + \cdots + \gamma_q)|u_1|^2 e^{-2\varphi}\, dV_\omega \\ &\leq \Big(\int_X (\gamma_1 + \cdots + \gamma_q)^{-1} |g|^2 e^{-2\varphi}\, dV_\omega \Big) \big(\|\overline{\nabla} u_1\|^2 + \|\overline{\nabla}^\star u_1\|^2 \big). \end{aligned}$$

Or $\overline{\nabla} u_1 = 0$ et $\overline{\nabla}^\star u_2 = 0$, de sorte que $\overline{\nabla}^\star u_1 = \overline{\nabla}^\star u$ et on a donc

$$\big|\langle\!\langle u, g \rangle\!\rangle\big|^2 \leq C \|\overline{\nabla}^\star u\|^2, \qquad C = \int_X (\gamma_1 + \cdots + \gamma_q)^{-1} |g|^2 e^{-2\varphi}\, dV_\omega.$$

La forme linéaire $v = \overline{\nabla}^\star u \mapsto \langle\!\langle u, g \rangle\!\rangle$ définie sur $\mathrm{Im}\, \nabla^\star$ se prolonge donc en une forme linéaire continue $v \mapsto \langle\!\langle v, f \rangle\!\rangle$ sur tout L^2, avec $\|f\|^2 \leq C$, et la relation $\langle\!\langle u, g \rangle\!\rangle = \langle\!\langle \overline{\nabla}^\star u, f \rangle\!\rangle$ implique par adjonction $\overline{\nabla} f = g$. Le théorème est donc démontré dans le cas d'une métrique h de classe C^∞. Dans le cas général, on utilise des arguments de régularisation. Comme la méthode est assez technique, nous renvoyons à [Dem82] pour les détails. L'idée est d'écrire $h = \lim_{\varepsilon \to 0} h_\varepsilon$, avec des poids φ_ε convergeant vers φ et formant une famille croissante par rapport à φ, en sorte que $e^{-\varphi_\varepsilon} \leq e^{-\varphi}$. On a donc $\|u\|_\varepsilon \leq \|u\|$ pour les normes L^2 associées. On trouve des solutions f_ε de l'équation $\overline{\nabla} f_\varepsilon = g$ satisfaisant l'estimation L^2

$$\begin{aligned} \int_X |f_\varepsilon|^2 e^{-2\varphi_\varepsilon}\, dV_\omega &\leq \int_X (\gamma_1 + \cdots + \gamma_q)^{-1} |g|^2 e^{-2\varphi_\varepsilon}\, dV_\omega \\ &\leq \int_X (\gamma_1 + \cdots + \gamma_q)^{-1} |g|^2 e^{-2\varphi}\, dV_\omega < +\infty. \end{aligned}$$

Le caractère uniforme de cette estimation vis-à-vis de ε entraîne qu'on peut extraire des f_ε une sous-suite convergeant en norme L^2 vers une solution f ayant les propriétés requises. Le point essentiel est de vérifier que le processus d'approximation est possible sans que les valeurs propres $\gamma_{j,\varepsilon}$ de $\Theta_{h_\varepsilon}(L) = \frac{i}{\pi}\partial\overline{\partial}\varphi_\varepsilon$ s'écartent trop de celles de $\frac{i}{\pi}\partial\overline{\partial}\varphi$. C'est effectivement possible d'après [Dem82, 92, 98]. Dans le cas où X est pseudoconvexe non compacte, le raisonnement est essentiellement identique, mais il faut se ramener d'abord au cas d'une métrique kählérienne complète pour éviter les difficultés éventuelles liées au domaine des opérateurs, qui pourraient introduire des "conditions au bord" non triviales. On s'appuie sur le fait qu'une variété kählérienne (X,ω) possédant une exhaustion plurisousharmonique $\psi \geq 0$ admet toujours une métrique kählérienne complète, par exemple $\widehat{\omega} = \omega + i\partial\overline{\partial}(\psi^2)$ (cf. [Dem82]).

2.4. Faisceaux d'idéaux multiplicateurs et théorème de Nadel

Nous introduisons maintenant le concept de *faisceau d'idéaux multiplicateurs*, suivant une définition donnée par A. Nadel [Nad89]. Nadel en a le premier donné une application marquante au problème de l'existence de métriques de Kähler-Einstein sur les variétés de Fano. L'idée de base remonte aux travaux fondamentaux de E. Bombieri [Bom70] et H. Skoda [Sko72a].

2.4.1. Définition. *Soit φ une fonction psh sur un ouvert $\Omega \subset X$; on associe à φ le faisceau d'idéaux $\mathcal{I}(\varphi) \subset \mathcal{O}_\Omega$, formé des germes de fonctions holomorphes $f \in \mathcal{O}_{\Omega,x}$ telles que $|f|^2 e^{-2\varphi}$ soit intégrable par rapport à la mesure de Lebesgue dans des coordonnées locales quelconques près de x. Ce faisceau sera appelé faisceau d'idéaux multiplicateurs associé au poids φ.*

La variété des zéros $V(\mathcal{I}(\varphi))$ est donc l'ensemble des points au voisinage desquels $e^{-2\varphi}$ est non intégrable. Bien entendu, de tel points ne peuvent apparaître que là où φ a des pôles logarithmiques. La formulation précise est la suivante.

2.4.2. Définition. *On dira qu'une fonction psh φ a un pôle logarithmique de coefficient $\gamma > 0$ en un point $x \in X$ si le nombre de Lelong*

$$\nu(\varphi, x) := \liminf_{z \to x} \frac{\varphi(z)}{\log|z-x|}$$

est égal à γ.

2.4.3. Lemme (Skoda [Sko72a]). *Soit φ une fonction psh sur un ouvert $\Omega \subset \mathbb{C}^n$ et soit $x \in \Omega$.*

(i) *Si* $\nu(\varphi,x)<1$, *alors* $e^{-2\varphi}$ *est intégrable au voisinage de* x, *en particulier on a* $\mathcal{I}(\varphi)_x=\mathcal{O}_{\Omega,x}$.

(ii) *Si* $\nu(\varphi,x)\geq n+s$ *pour un certain entier* $s\geq 0$, *alors* $e^{-2\varphi}\geq C|z-x|^{-2n-2s}$ *au voisinage de* x *et* $\mathcal{I}(\varphi)_x\subset\mathfrak{m}_{\Omega,x}^{s+1}$, *où* $\mathfrak{m}_{\Omega,x}$ *désigne l'idéal maximal de* $\mathcal{O}_{\Omega,x}$.

Preuve. La démonstration repose sur des estimations classiques de théorie du potentiel complexe, voir H. Skoda [Sko72a].

2.4.4. Proposition ([Nad89]). *Pour toute fonction psh* φ *sur* $\Omega\subset X$, *le faisceau* $\mathcal{I}(\varphi)$ *est un faisceau cohérent d'idéaux sur* Ω.

Preuve. Puisque le résultat est local, nous pouvons supposer que Ω est la boule unité de $\mathbb{C}^n$. Soit $\mathcal{H}_\varphi(\Omega)$ l'ensemble des fonctions f holomorphes sur Ω telles que $\int_\Omega|f|^2e^{-2\varphi}dV<+\infty$. D'après la propriété noethérienne forte des faisceaux cohérents, l'ensemble $\mathcal{H}_\varphi(\Omega)$ engendre un faisceau d'idéaux cohérent $\mathcal{J}\subset\mathcal{O}_\Omega$. Il est clair que $\mathcal{J}\subset\mathcal{I}(\varphi)$; pour démontrer l'égalité, il suffit de vérifier que $\mathcal{J}_x+\mathcal{I}(\varphi)_x\cap\mathfrak{m}_{\Omega,x}^{s+1}=\mathcal{I}(\varphi)_x$ pour tout entier s, en vertu du lemme de Krull. Soit $f\in\mathcal{I}(\varphi)_x$ un germe défini sur un voisinage V de x et soit θ une fonction tronquante à support dans V, telle que $\theta=1$ au voisinage de x. On résout l'équation $\overline{\partial}u=g:=\overline{\partial}(\theta f)$ au moyen des estimations L^2 de Hörmander (14.3), où L est le fibré en droites trivial $\Omega\times\mathbb{C}$ muni du poids strictement psh

$$\widetilde{\varphi}(z)=\varphi(z)+(n+s)\log|z-x|+|z|^2.$$

Nous obtenons une solution u telle que $\int_\Omega|u|^2e^{-2\varphi}|z-x|^{-2(n+s)}dV<\infty$, donc $F=\theta f-u$ est holomorphe, $F\in\mathcal{H}_\varphi(\Omega)$ et $f_x-F_x=u_x\in\mathcal{I}(\varphi)_x\cap\mathfrak{m}_{\Omega,x}^{s+1}$. Ceci démontre notre affirmation.

2.4.5. Théorème d'annulation de Nadel ([Nad89], [Dem93]). *Soit* (X,ω) *une variété kählérienne compacte* (*ou faiblement pseudoconvexe*), *et soit* L *un fibré en droites holomorphe sur* X *muni d'une métrique hermitienne* h *singulière de poids* φ *telle que* $\Theta_h(L)\geq\varepsilon\omega$, $\varepsilon>0$. *On note* $\mathcal{I}(h)=\mathcal{I}(\varphi)$. *Alors*

$$H^q\big(X,\mathcal{O}(K_X+L)\otimes\mathcal{I}(h)\big)=0\qquad\text{pour tout } q\geq 1.$$

Preuve. Soit $\mathcal{L}^q$ le faisceau des germes de (n,q)-formes u à valeurs dans L et à coefficients mesurables, telles que à la fois $|u|^2e^{-2\varphi}$ et $|\overline{\partial}u|^2e^{-2\varphi}$ soient localement intégrables. L'opérateur $\overline{\partial}$ définit un complexe de faisceaux $(\mathcal{L}^\bullet,\overline{\partial})$ qui est une résolution du faisceau $\mathcal{O}(K_X+L)\otimes\mathcal{I}(\varphi)$: en effet, le noyau de $\overline{\partial}$ en degré 0 consiste en les germes de n-formes holomorphes à valeurs dans L qui satisfont la condition d'intégrabilité; donc la fonction coefficient appartient à $\mathcal{I}(\varphi)$; l'exactitude en degré $q\geq 1$ découle du

Théorème 2.3.1 appliqué à des boules arbitrairement petites. Comme chaque faisceau $\mathcal{L}^q$ est un $\mathcal{C}^\infty$-module, $\mathcal{L}^\bullet$ est une résolution par des faisceaux acycliques. En appliquant maintenant le théorème 2.3.1 globalement sur X, on déduit de l'isomorphisme de De Rham-Weil que $H^q\big(\Gamma(X,\mathcal{L}^\bullet)\big) = 0$ pour $q \geq 1$ [si X est non compacte, on choisit une fonction d'exhaustion plurisousharmonique ψ de classe C^∞ sur X, et on multiplie la métrique initiale de L par le facteur $e^{-\chi\circ\psi}$, où χ est une fonction convexe croissante de croissance arbitrairement rapide à l'infini ; ceci permet d'assurer la convergence des intégrales à l'infini]. Le théorème est démontré.

2.4.6. Corollaire. *Soient (X,ω), L et φ comme dans le théorème* 2.4.5 *et soient $x_1, \ldots, x_N$ des points isolés de la variété des zéros $V(\mathcal{I}(\varphi))$. Alors il existe une application surjective*

$$H^0(X,\mathcal{O}(K_X+L)) \longrightarrow\!\!\!\!\!\rightarrow \bigoplus_{1\leq j\leq N} \mathcal{O}(K_X+L)_{x_j} \otimes \big(\mathcal{O}_X/\mathcal{I}(\varphi)\big)_{x_j}.$$

Preuve. Considérons la suite exacte longue de cohomologie associée à la suite exacte courte $0 \to \mathcal{I}(\varphi) \to \mathcal{O}_X \to \mathcal{O}_X/\mathcal{I}(\varphi) \to 0$, tordue par $\mathcal{O}(K_X+L)$, et appliquons le théorème 2.4.5 pour obtenir l'annulation du premier groupe H^1. La propriété de surjectivité annoncée s'ensuit.

2.4.7. Corollaire. *Soient (X,ω), L et φ comme dans le théorème* 2.4.5. *Supposons que la fonction poids φ soit telle que $\nu(\varphi,x) \geq n+s$ en un certain point $x \in X$, et $\nu(\varphi,y) < 1$ pour $y \neq x$ assez voisin de x. Alors $H^0(X,K_X+L)$ engendre tous les s-jets de sections au point x.*

Preuve. Le lemme de Skoda 2.4.3 (ii) montre que $e^{-2\varphi}$ est intégrable au voisinage de tout point $y \neq x$ suffisamment proche de x, donc $\mathcal{I}(\varphi)_y = \mathcal{O}_{X,y}$, alors que $\mathcal{I}(\varphi)_x \subset \mathfrak{m}_{X,x}^{s+1}$ d'après 2.4.3 (ii). Le corollaire 2.4.7 est donc un cas particulier de 2.4.6.

2.4.8. Commentaire. La philosophie générale de ces résultats (qui peuvent être considérés comme des généralisations du théorème de Hörmander-Bombieri-Skoda [Bom70], [Sko72a, 75]) est la suivante: le problème de construire des sections holomorphes de K_X+L peut se résoudre en construisant des métriques singulières convenables sur L, telles que le poids φ ait des pôles logarithmiques en des points donnés x_j, ces pôles étant isolés ou entourés de basses multiplicités avoisinantes.

2.4.9. Cas particulier. Si X est compacte et si (L,h_0) est un fibré en droites hermitien C^∞ à courbure positive sur X, on va pouvoir construire une métrique singulière h ayant des pôles logarithmiques isolés en un nombre fini de points $\{x_1, \ldots, x_N\}$ quelconques. Il suffit de poser $h = h_0 e^{-\psi}$ avec $\psi(z) = \varepsilon \sum \theta_j(z^{(j)}) \log|z^{(j)}|$, où les $z^{(j)}$

sont des coordonnées locales centrées en x_j et les θ_j des fonctions plateau adéquates. La courbure de h restera positive si ε est assez petit. On voit alors qu'il existe des constantes $a, b \geq 0$ (dépendant seulement de L et N) telles que pour tout $s \in \mathbb{N}$ le groupe $H^0(X, \mathcal{O}(mL))$ engendre les jets d'ordre s en tout point x_j pour $m \geq as + b$. Ceci permet de démontrer le théorème de plongement de Kodaira (cf. § 2.5). On peut aussi de la même manière résoudre le problème de Levi (à savoir, montrer qu'une variété fortement pseudoconvexe est de Stein): il suffit de travailler sur le fibré trivial $L = \mathcal{O}_X$, muni de poids $e^{-\psi}$ où ψ est exhaustive, fortement plurisousharmonique et suffisamment grande.

2.4.10. Remarque. Le théorème de Nadel contient en fait un théorème d'annulation fondamental de la géométrie algébrique, démontré indépendamment par Kawamata [Kaw82] et Viehweg [Vie82]. Le théorème de Kawamata-Viehweg, au changement de vocabulaire près, correspond au cas particulier où X est algébrique projective, la métrique singulière h présentant des pôles logarithmiques le long d'un diviseur à croisement normaux (on peut ensuite se ramener au cas d'un diviseur quelconque par le théorème de désingularisation de Hironaka).

2.5. Diverses notions de fibrés (semi-)positifs, amples, nefs, gros, etc...

Nous rappelons ici quelques notions liées à la positivité, et qui sont particulièrement utiles en géométrie algébrique. Si $Y \subset X$ est une sous-variété de dimension d, on notera comme d'habitude

$$L^d \cdot Y = \int_Y c_1(L)^d.$$

2.5.1. Définition. *Un fibré en droites L sur une variété projective X est dit*

(i) *engendré par ses sections si le morphisme de restriction $H^0(X, L) \to L_x$ est surjectif en tout point $x \in X$, ce qui revient à dire que l'ensemble base $B_{|L|}$ du système linéaire complet $|L| = P(H^0(X, L))$ est vide;*

(ii) *très ample si $B_{|L|} = \emptyset$ et si l'application $\Phi_{|L|} : X \to \mathbb{P}(H^0(X, L)^\star)$ est un plongement;*

(iii) *semi-ample s'il existe un multiple mL, $m > 0$, qui soit engendré par ses sections;*

(iv) *ample s'il existe un multiple mL, $m > 0$, qui soit très ample;*

(v) *nef, si $L \cdot C \geq 0$ pour toute courbe algébrique C;*

(vi) *effectif, si mL possède une section pour au moins un $m > 0$;*

(vii) *pseudo-effectif, si la classe de Chern $c_1(L)$ appartient au cône fermé de $H^{1,1}(X)$ engendré par les classes de diviseurs effectifs*;

(viii) *gros, si la "dimension de Kodaira-Iitaka"*

$$\kappa(L) = \limsup_{m\to+\infty} \frac{\log \dim H^0(X, mL)}{\log m} \in \{-\infty, 0, 1, \ldots, n\},$$

est égale à $n = \dim X$.

Certaines de ces notions ne dépendent que de la classe de Chern de L (on dit que ce sont des notions "numériques"): ce sont les notions de fibré ample, nef, pseudo-effectif et gros. Pour ces quatre notions, il y a en fait un dictionnaire algébro-analytique qui donne des traductions en termes de propriétés de la courbure (cf. par exemple [Dem90]).

2.5.2. Proposition. *Soit X une variété projective munie d'une métrique hermitienne ω. Un fibré L sur X est*

(i) *ample, si et seulement si L possède une métrique hermitienne C^∞ à courbure positive (on dit alors aussi que L est positif)*;

(ii) *nef, si et seulement si pour tout $\varepsilon > 0$, il existe une métrique hermitienne h_ε de classe C^∞ telle que $\Theta_{h_\varepsilon}(L) \geq -\varepsilon\omega$*;

(iii) *pseudo-effectif, si et seulement si L possède une métrique singulière h telle que $\Theta_h(L) \geq 0$*;

(iv) *gros, si et seulement si L possède une métrique singulière h telle que $\Theta_h(L) \geq \varepsilon\omega$ pour un certain $\varepsilon > 0$.*

2.5.3. Commentaires. L'assertion (i) n'est autre que le célèbre théorème de plongement de Kodaira. En effet, si mL est engendré par ses sections, l'exemple 2.1.6 montre que $\Phi = \Phi_{|mL|}$ induit une métrique à courbure semi-positive sur $mL = \Phi^*\mathcal{O}(1)$ (et donc aussi sur L en prenant une racine m-ième de la métrique); la courbure est bien définie positive si $\Phi_{|mL|}$ est un plongement. Dans la direction inverse, si (L, h) est C^∞ à courbure positive, 2.4.9 entraîne que mL est très ample pour m assez grand. Si h est une métrique singulière à courbure définie positive $\Theta_h(L) \geq \varepsilon\omega$, le Corollaire 2.4.7 montre que l'on peut construire des sections de mL ayant des jets prescrits, mais seulement aux points de x où le poids φ n'a pas trop de singularités (par exemple, là où $\nu(\varphi, x) = 0$); ceci implique alors que L est gros, et la réciproque résulte assez facilement de 2.1.6.

Toutes ces notions sont sans doute mieux appréhendées si on introduit le cône nef K_{nef}, i.e. le cône fermé de $H^{1,1}(X)$ engendré par les classes de diviseurs nef, et

de manière analogue, le cône pseudo-effectif $K_{\mathrm{psef}} \subset H^{1,1}(X)$. Alors le cône "ample" est précisément l'intérieur de K_{nef}, et le cône "gros" l'intérieur de K_{psef}. On appelle par définition *dimension de Kodaira* de la variété X l'entier

$$\kappa(X) := \kappa(K_X) \in \{-\infty, 0, 1, \ldots, n\},$$

et on dit que X est de *type général* si K_X est gros, i.e. si $\kappa(X) = n$.

3. APPLICATION À LA CONJECTURE DE FUJITA

3.1. Méthode de l'équation de Monge-Ampère

En dehors du cas des dimensions 1 et 2, le premier résultat général sur la conjecture de Fujita est celui de [Dem93] (le manuscrit date de 1991). La méthode développée dans ce travail consistait à produire des métriques singulières sur L en résolvant une équation de Monge-Ampère de la forme

$$\Big(\Theta_h(L) + \frac{i}{\pi}\partial\overline{\partial}\psi_\varepsilon\Big)^n = f_\varepsilon,$$

et en laissant le second membre converger vers une combinaison linéaire de mesures de Dirac $\sum \rho_j \delta_{x_j}$ – la masse totale étant ajustée à la valeur requise $c_1(L)^n$ par ajout d'une forme volume à densité C^∞. De cette façon, il est possible d'obtenir à la limite des métriques singulières $he^{-2\psi}$, $\psi = \lim \psi_\varepsilon$, présentant des singularités logarithmiques aux points x_j. Le contrôle des nombres de Lelong aux points $y \neq x_j$ se fait en ajustant précisément le choix des constantes ρ_j et utilisant la théorie de l'intersection des courants. Nous ne développerons pas ici cette méthode, car les résultats effectifs obtenus sont assez nettement moins bons que ceux obtenus par les méthodes ultérieures.

3.2. Utilisation de la formule de Riemann-Roch

À la suite des premiers résultats de [Dem93], de nombreux travaux se sont attachés à améliorer les bornes effectives obtenues et à développer des approches plus algébriques. Kollár [Kol92] obtient ainsi une preuve purement algébrique dans l'esprit du "théorème de non-annulation" de Shokurov. Ein-Lazarsfeld [EL93] et Fujita [Fuj93] ont établi la partie (i) de la conjecture de Fujita en dimension 3, et un raffinement très poussé de leur technique a permis à Kawamata [Kaw95] d'atteindre le cas de la dimension 4 (Helmke [Hel96] a ensuite simplifié cette approche, et vient

d'annoncer le cas de la dimension 5). Dans cette section, nous présentons une approche algébrique relativement élémentaire due à Y.T. Siu [Siu96]. L'idée est d'utiliser la formule de Riemann-Roch :

3.2.1. Cas particulier de la formule de Riemann-Roch. *Soit $\mathcal{J} \subset \mathcal{O}_X$ un faisceau d'idéaux cohérent sur X tel que la variété des zéros $V(\mathcal{J})$ soit de dimension d (avec éventuellement des composantes de dimension plus basse). Soit $Y = \sum \lambda_j Y_j$ le cycle algébrique effectif de dimension d associé aux composantes de dimension d de $V(\mathcal{J})$ (les multiplicités λ_j prennent en compte les multiplicités de l'idéal $\mathcal{J}$ le long de chaque composante). Alors, pour tout fibré en droites L, la caractéristique d'Euler $\chi(X, \mathcal{O}(E + mL) \otimes \mathcal{O}_X/\mathcal{O}(\mathcal{J}))$ est un polynôme $P(m)$ de degré d et de coefficient directeur $L^d \cdot Y/d!$*

À l'aide de ce lemme, on démontre simultanément les deux théorèmes suivants.

3.2.2. Théorème (Fujita). *Si L est un fibré en droites ample sur une variété projective X de dimension n, alors $K_X + (n+1)L$ est nef.*

3.2.3. Théorème (Siu). *Soit L comme ci-dessus et soit G un fibré en droites nef. On a les propriétés suivantes.*

(i) *$2K_X + mL + G$ engendre les jets simultanés d'ordre $s_1, \ldots, s_p \in \mathbb{N}$ en des points arbitraires $x_1, \ldots, x_p \in X$, i.e., il existe une application surjective*

$$H^0(X, \mathcal{O}(2K_X + mL + G)) \longrightarrow\!\!\!\!\!\rightarrow \bigoplus_{1 \le j \le p} \mathcal{O}(2K_X + mL + G) \otimes \mathcal{O}_{X,x_j}/\mathfrak{m}_{X,x_j}^{s_j+1},$$

dès que $m \ge 2 + \sum_{1 \le j \le p} \binom{3n + 2s_j - 1}{n}$.

En particulier $2K_X + mL + G$ est très ample pour $m \ge 2 + \binom{3n+1}{n}$.

(ii) *$2K_X + (n+1)L + G$ engendre les jets simultanés d'ordre $s_1, \ldots, s_p$ en des points arbitraires $x_1, \ldots, x_p \in X$ dès que les nombres d'intersection $L^d \cdot Y$ de L sur tous les sous-ensembles algébriques Y de X de dimension d sont tels que*

$$L^d \cdot Y > \frac{2^{d-1}}{\lfloor n/d \rfloor^d} \sum_{1 \le j \le p} \binom{3n + 2s_j - 1}{n}.$$

Schéma de la preuve. Les démonstrations de 3.2.2 et 3.2.3 (i, ii) sont tout à fait parallèles, nous renvoyons à [Siu96] et [Dem96] pour les détails. L'idée est de trouver un

entier (ou un rationnel) m_0 et une métrique hermitienne singulière h_0 sur $K_X + m_0 L$ dont le courant de courbure est strictement positif, $\Theta_{h_0} \geq \varepsilon\omega$, telle que $V(\mathcal{I}(h_0))$ soit de dimension 0 et telle que le poids φ_0 de h_0 satisfasse $\nu(\varphi_0, x_j) \geq n + s_j$ pour tout j [dans le cas de 3.2.2, on convient simplement que $\{x_1, \ldots, x_p\} = \emptyset$, et on cherche à obtenir m_0 arbitrairement proche de $n+1$]. Comme L et G sont nefs, 2.5.2 (ii) implique que $(m - m_0)L + G$ possède pour tout $m \geq m_0$ une métrique h' dont la courbure $\Theta_{h'}$ a une partie négative arbitrairement petite, disons $\Theta_{h'} \geq -\frac{\varepsilon}{2}\omega$. Alors $\Theta_{h_0} + \Theta_{h'} \geq \frac{\varepsilon}{2}\omega$ est positive définie. Une application du Cor. 2.4.7 au fibré $F = K_X + mL + G = (K_X + m_0 L) + ((m - m_0)L + G)$ muni de la métrique $h_0 \otimes h'$ garantit l'existence des sections de $K_X + F = 2K_X + mL + G$ réalisant les jets désirés pour $m \geq m_0$.

Fixons un plongement $\Phi_{|\mu L|} : X \to \mathbb{P}^N$, $\mu \gg 0$, donné par des sections $\lambda_0, \ldots, \lambda_N$ de $H^0(X, \mu L)$, et soit h_L la métrique associée sur L, de forme de courbure définie positive $\omega = \Theta(L)$. Pour obtenir la métrique désirée h_0 sur $K_X + m_0 L$, on fixe un entier $a \in \mathbb{N}^*$ et on utilise un procédé de récurrence double pour construire des métriques singulières $(h_{k,\nu})_{\nu \geq 1}$ sur $aK_X + b_k L$, pour une suite décroissante (au sens large) d'entiers positifs $b_1 \geq b_2 \geq \cdots \geq b_k \geq \cdots$. Une telle suite est nécessairement stationnaire et m_0 sera précisément la limite stationnaire $m_0 = \lim b_k/a$. L'idée principale qui fait fonctionner la récurrence réside dans le fait qu'une "bonne" métrique sur $L' = aK_X + b_k L$ permet de construire des sections de $K_X + L' = (a+1)K_X + b_k L$, qui à leur tour fournissent des métriques sur $\frac{a}{a+1}(K_X + L') = aK_X + \frac{a}{a+1} b_k L$. En fait, il se produit une petite perte sur le multiple de L due au fait qu'on va appliquer la formule de Riemann-Roch et qu'il faut éviter les zéros et les petites valeurs du polynôme de Hilbert. On trouve ainsi une estimation $b_{k+1} \leq \frac{a}{a+1} b_k + N$ où N est un entier explicite. De cette façon, on réduit petit à petit la taille de l'entier b_k, en partant de n'importe quel entier b_1 suffisamment grand pour initialiser la récurrence (et qu'on ne cherche pas à contrôler). Les métriques $h_{k,\nu}$ sont choisies en sorte qu'elles satisfassent les propriétés suivantes :

(α) $h_{k,\nu}$ est une métrique "algébrique" de la forme

$$\|\xi\|^2_{h_{k,\nu}} = \frac{|\tau_k(\xi)|^2}{\big(\sum_{1 \leq i \leq \nu,\, 0 \leq j \leq N} \big|\tau_k^{(a+1)\mu}(\sigma_i^{a\mu} \cdot \lambda_j^{(a+1)b_k - am_i})\big|^2\big)^{1/(a+1)\mu}},$$

définie par des sections $\sigma_i \in H^0\big(X, \mathcal{O}((a+1)K_X + m_i L)\big)$, $m_i < \frac{a+1}{a} b_k$, $1 \leq i \leq \nu$, où $\xi \mapsto \tau_k(\xi)$ est une trivialisation locale arbitraire de $aK_X + b_k L$;

notons que $\sigma_i^{a\mu} \cdot \lambda_j^{(a+1)b_k - am_i}$ est une section de

$$a\mu((a+1)K_X + m_i L) + ((a+1)b_k - am_i)\mu L = (a+1)\mu(aK_X + b_k L).$$

(β) $\mathrm{ord}_{x_j}(\sigma_i) \geq (a+1)(n+s_j)$ pour tout i, j ;

(γ) $\mathcal{I}(h_{k,\nu+1}) \supset \mathcal{I}(h_{k,\nu})$ et $\mathcal{I}(h_{k,\nu+1}) \neq \mathcal{I}(h_{k,\nu})$ tant que la variété des zéros $V(\mathcal{I}(h_{k,\nu}))$ est de dimension positive.

La métrique $h_{k,\nu+1}$ est produite à partir de $h_{k,\nu}$ en appliquant la formule de Riemann-Roch 3.2.1 sur $Y = Y_{k,\nu} = V(\mathcal{I}(h_{k,\nu}))$. On obtient beaucoup de sections sur Y, et on peut donc leur imposer la multiplicité de zéros voulue en comparant le nombre de sections au nombre de conditions linéaires requises. Par ailleurs, ces sections s'étendent à X car le théorème de Nadel donne précisément l'annulation du H^1 à valeurs dans l'idéal $\mathcal{I}_{h_{k,\nu}}$. On déduit alors $h_{k,\nu+1}$ de $h_{k,\nu}$ par rajout d'une section non triviale sur Y. Le procédé s'arrête nécessairement à $\dim Y_{k,\nu} = 0$ d'après (γ), parce qu'une suite décroissante d'ensembles algébriques est toujours stationnaire par le théorème de Noether. Il resterait à estimer précisément les entiers b_k successifs produits par la formule de Riemann-Roch. Nous renvoyons à [Dem96] pour le détail des calculs. □

3.3. Méthode de Angehrn-Siu

Il existe une autre méthode que celle décrite à la section 3.2, permettant d'obtenir des métriques singulières h sur L ayant la propriété que la variété $V(\mathcal{I}(h))$ soit constituée de point isolés, et produisant des bornes effectives substantiellement meilleures. Cette méthode est due à Angehrn-Siu [AS95] (cf. aussi H. Tsuji [Tsu96] pour une approche voisine). L'idée consiste à partir d'une métrique h_∞ de classe C^∞ à courbure positive, et de prendre une métrique h_1 singulière ayant des singularités assez fortes en un point $a \in X$ prescrit. Si l'on veut que $V(\mathcal{I}(h_1))$ contienne a, il suffit que le poids φ_1 de h_1 satisfasse $\nu(\varphi_1, a) \geq \alpha n$, $\alpha \geq 1$, et pour cela il suffit de poser $\varphi_1 = \frac{1}{m}\log|\sigma_1|$ avec une section $\sigma_1 \in H^0(X, mL)$ ayant en a une multiplicité $> \alpha mn$. D'après Riemann-Roch, c'est possible pour m assez grand dès lors que $L^n > \alpha^n n^n$. Il se peut malheureusement que a ne soit pas isolé dans $V(\mathcal{I}(h_1))$. Ce que l'on fait alors est de remplacer h_1 par une interpolation linéaire $h_t = h_1^t \cdot h_\infty^{1-t}$ [de poids $\psi_t = t\varphi_1 + t\varphi_\infty$] avec $t = \tau_1 - \delta_1$ très légèrement inférieur à l'inf τ_1 des valeurs t pour lesquelles $Y_{\tau_1} = Y_t = V(\mathcal{I}(h_t))$ a une composante de dimension positive passant par a ; par construction $\tau_1 \leq \frac{1}{\alpha}$. Alors $V(\mathcal{I}(h_{\tau_1-\delta_1}))$ ne contient peut être plus a, mais il suffira de remplacer le poids lisse φ_∞ par un poids ayant une certaine singularité au poids a pour que a figure de nouveau dans la variété $V(\mathcal{I}(\psi_t))$ ainsi obtenue. De façon précise, on remplace $\psi_t^{(1)} := \psi_t$ par

$$\psi_t^{(2)} = (\tau_1 - \delta_1))\varphi_1 + t\varphi_2 + (1 - \tau_1 - t + \delta_1)\varphi_\infty$$

où $\varphi_{2|Y_{\tau_1}}$ n'est pas identiquement $-\infty$, a une forte singularité en a (tout en étant peut-être très peu singulière au point générique de Y_{τ_1}). Ce poids s'obtient en construisant une section $\sigma_2 \in H^0(Y_{\tau_1}, pL)$ s'annulant fortement en a (on utilise de nouveau Riemann-Roch), et en étendant une certaine puissance σ_2^q (non contrôlée a priori) en une section $\widetilde{\sigma}_2$ de X ; on pose alors $\varphi_2 = \frac{1}{pq}\log|\widetilde{\sigma}_2|$. Par construction, pour δ_1 assez petit et $t = 1$, le poids $\psi_t^{(2)}$ est tel que $\dim V(\mathcal{I}(\psi_t^{(2)})) < \dim Y_{\tau_1}$. On prend pour τ_2 l'inf des t pour lesquels $a \in V(\mathcal{I}(\psi_t^{(2)}))$ et on fixe $t = \tau_2 - \delta_2$ très légèrement inférieur à τ_2. On procède ainsi par récurrence de façon à obtenir une suite emboîtée d'ensembles algébriques

$$Y_{\tau_p}^{(p)} \subset \cdots \subset Y_{\tau_1}^{(1)}$$

contenant a, dont les dimensions au voisinage de a sont strictement décroissantes. On a "gagné" lorsque $\dim Y_{\tau_p}^{(p)} = 0$. En réalité, il se présente de multiples difficultés dues au fait que les ensembles $Y_{\tau_p}^{(p)}$ ne sont ni lisses ni irréductibles. On surmonte ces difficultés en effectuant des résolutions de singularités et en perturbant un peu les multiplicités de façon à faire apparaître une seule composante utile à chaque étape (méthode de Shokurov [Sho85]). On utilise aussi des arguments de semi-continuité pour les idéaux multiplicateurs, à partir de calculs faits aux points génériques (voir [Kol95], [DK99] pour des versions élaborées des théorèmes de semi-continuité). En effectuant un comptage précis des multiplicités, on aboutit aux résultats effectifs suivants dûs à Angehrn-Siu [AS95] (voir aussi [Tsu96]).

3.3.1. Théorème. *Si $m \geq \frac{1}{2}(n^2+n+2)$, alors $K_X + mL$ est engendré par ses sections globales.*

3.3.2. Théorème. *Si $(L^d \cdot Y)^{\frac{1}{d}} > \frac{1}{2}n(n+2r-1)$ pour toute sous-variété irréductible $Y \subset X$ de dimension $d > 0$, alors les sections holomorphes globales de $K_X + L$ séparent r points distincts quelconques $x_1, \ldots, x_r$ de X.*

Il est à noter toutefois qu'on ne peut pas atteindre les jets d'ordre 1 et plus par cette méthode, car on s'astreint à réaliser les plus petites singularités possibles qui fassent apparaître des idéaux de Nadel non triviaux.

3.4. Version effective du grand théorème de Matsusaka

Il s'agit du résultat suivant, essentiellement dû à Y.T. Siu [Siu93], donnant une version effective des théorèmes de Matsusaka [Mat72] et Kollár-Matsusaka [KoM83].

3.4.1. Théorème. *Soient L et B des fibrés en droites nefs sur une variété X projective de dimension n. Supposons que L soit ample et soit $H = n^3(K_X + (n+2)L)$.*

Alors $mL - B$ est très ample pour

$$m \geq (2n)^{(3^{n-1}-1)/2} \frac{(L^{n-1} \cdot (B+H))^{(3^{n-1}+1)/2} (L^{n-1} \cdot H)^{3^{n-2}(n/2-3/4)-1/4}}{(L^n)^{3^{n-2}(n/2-1/4)+1/4}}.$$

En particulier mL est très ample pour

$$m \geq C_n \, (L^n)^{3^{n-2}} \left(n + 2 + \frac{L^{n-1} \cdot K_X}{L^n} \right)^{3^{n-2}(n/2+3/4)+1/4}$$

avec $C_n = 2^{(3^{n-1}-1)/2} n^{3^{n-1}(n/2+5/4)-1/4}$.

Pour les surfaces ($n = 2$), la méthode fournit la borne

$$m \geq 4 \frac{(L \cdot (K_X + 4L))^2}{L^2}. \tag{3.4.2}$$

Ceci est pratiquement optimal d'après les travaux de Fernández del Busto [FdB95, 96]. Nous ne dirons pas grand chose de la preuve, sauf que l'on utilise de manière cruciale le théorème d'annulation de Nadel, les résultats d'amplitude effective des sections 3.2 et 3.3, et un cas particulier des inégalités de Morse [Dem85].

4. INVARIANCE DES PLURIGENRES PAR DÉFORMATION

Étant donné une déformation $\gamma : X \to S$ de variétés projectives, le problème de l'invariance des plurigenres des fibres $X_t = \gamma^{-1}(t)$ se ramène au cas où la base est le disque unité $\Delta \subset \mathbb{C}$ (en général, on peut toujours relier deux points quelconques de S par une chaîne de petits disques analytiques, et il suffit de restreindre la base à chacun de ces disques). On suppose donc $\gamma : X \to \Delta$. Dans ce cas, on a un isomorphisme canonique $K_{X_t} \simeq K_{X|X_t}$ sur chaque fibre, donné par $u \mapsto dt \wedge u$ [de sorte qu'on se permettra d'identifier K_{X_t} et $K_{X|X_t}$ dans la suite]. On sait d'après le théorème des images directes de Grauert [Gra60] que les faisceaux images directes $\gamma_\star \mathcal{O}(mK_X)$ sont cohérents, et de plus, les plurigenres $p_m(X_t) = h^0(X_t, (mK_X)_{|X_t})$ sont des fonctions semi-continues supérieurement de t. Les sauts se produisent précisément si une section sur une certaine fibre X_{t_0} ne se prolonge pas aux fibres voisines. Démontrer l'invariance des plurigenres revient donc à montrer qu'une section de $mK_{X|X_{t_0}}$ sur une fibre X_{t_0} se prolonge sur un voisinage de X_{t_0} dans X. On pourra supposer sans perte de généralité que $t_0 = 0$. La stratégie consiste grosso modo à construire des

métriques singulières sur K_X et K_{X_0} de la façon "la plus canonique possible", puis à utiliser des théorèmes d'extensions L^2 par rapport à ces métriques.

4.1. Métriques à singularités minimales

L'une des idées essentielles de la démonstration du théorème d'invariance des plurigenres – bien qu'elle ne soit pas tout à fait explicite dans [Siu97] – réside dans le fait que les singularités des métriques hermitiennes sur un fibré L reflètent de manière très intime les ensembles-base des systèmes linéaires $|mL|$, pourvu qu'on choisisse précisément les métriques à singularités minimales. Nous suivons ici l'approche de [DPS98].

4.1.1. Définition. *Soit L un fibré en droites pseudo-effectif sur une variété complexe compacte X. Considérons des métriques hermitiennes h_1, h_2 sur L à courbure $\Theta_{h_j}(L) \geq 0$ au sens des courants.*

(i) *On écrira $h_1 \preccurlyeq h_2$, et on dira que h_1 est moins singulière que h_2, s'il existe une constante $C > 0$ telle que $h_1 \leq Ch_2$.*

(ii) *On écrira $h_1 \sim h_2$, et on dira que h_1, h_2 sont équivalentes du point de vue des singularités, s'il existe une constante $C > 0$ telle que $C^{-1}h_2 \leq h_1 \leq Ch_2$.*

Bien entendu $h_1 \preccurlyeq h_2$ si et seulement si les poids associés dans des trivialisations vérifient localement $\varphi_2 \leq \varphi_1 + C$, ce qui implique en particulier $\nu(\varphi_1, x) \leq \nu(\varphi_2, x)$ en tout point. La définition ci-dessus est motivée par l'observation suivante.

4.1.2. Théorème. *Pour tout fibré en droites L pseudo-effectif au-dessus d'une variété X complexe compacte, il existe à équivalence près de singularités une unique classe de métriques hermitiennes h à singularités minimales telles que $\Theta_h(L) \geq 0$.*

Preuve. C'est quasiment trivial. On fixe une fois pour toutes une métrique h_∞ de classe C^∞ (dont la courbure est de signature arbitraire variable), et on écrit les métriques singulières de L sous la forme $h = h_\infty e^{-2\psi}$. La condition $\Theta_h(L) \geq 0$ équivaut à supposer $\frac{i}{\pi}\partial\overline{\partial}\psi \geq -u$ où $u = \Theta_{h_\infty}(L)$. Cette condition entraîne que ψ est plurisousharmonique à l'ajout près du poids φ_∞ de h_∞, et donc localement majorée. Comme on s'intéresse aux métriques uniquement à équivalence de singularités près, on peut toujours ajuster ψ par une constante en sorte que $\sup_X \psi = 0$. On pose maintenant

$$h_{\min} = h_\infty e^{-2\psi_{\min}}, \qquad \psi_{\min}(x) = \sup_\psi \psi(x)$$

où le sup est étendu à toutes les fonctions ψ telles que $\sup_X \psi = 0$ et $\frac{i}{\pi}\partial\overline{\partial}\psi \geq -u$. D'après les résultats standards sur les fonctions plurisousharmoniques (cf. Lelong

[Lel69]), $\psi_{\min}$ vérifie encore $\frac{i}{\pi}\partial\overline{\partial}\psi_{\min} \geq -u$ (i.e. le poids $\varphi_\infty + \psi_{\min}$ de $h_{\min}$ est plurisousharmonique), et $h_{\min}$ est évidemment la métrique à singularité minimale cherchée. □

Étant donné une section $\sigma \in H^0(X, mL)$ alors $h(\xi) = |\xi^m/\sigma(x)|^{1/m}$ définit évidemment une métrique singulière sur L, qui a nécessairement au moins autant de singularités que $h_{\min}$, i.e. $\frac{1}{m}\log|\sigma| \leq \varphi_{\min} + C$ localement. En particulier $|\sigma|^2 e^{-2m\varphi_{\min}}$ est localement bornée, en sorte que $\sigma \in H^0(X, mL \otimes \mathcal{I}(h_{\min}^m))$, et on a donc pour tout $m > 0$ un isomorphisme

$$H^0(X, mL \otimes \mathcal{I}(h_{\min}^m)) \xrightarrow{\simeq} H^0(X, mL).$$

En vertu de cet isomorphisme, on dira que l'ensemble

$$E_+(h_{\min}) = \{x \in X\,;\ \nu(\varphi_{\min}, x) > 0\} \tag{4.1.3}$$

est *l'ensemble-base virtuel* de L. On a toujours $E_+(h_{\min}) \subset \bigcap_{m>0} B_{|mL|}$ et il peut y avoir inclusion stricte, notamment si les mL n'ont aucune section non nulle (c'est le cas si L est un fibré plat générique, bien qu'on ait alors $E_+(h_{\min}) = \emptyset$). Si h est une métrique hermitienne singulière telle que $\Theta_h(L) \geq 0$ et

$$H^0(X, mL \otimes \mathcal{I}(h^m)) \simeq H^0(X, mL) \qquad \text{pour tout } m \geq 0, \tag{4.1.4}$$

on dira que h est une *décomposition de Zariski analytique* de L. On vient de voir qu'une telle décomposition existe toujours et que $h = h_{\min}$ répond à la question. Cette définition est motivée par l'analogue algébrique (qui ne conduit pas toujours à une réponse affirmative, loin s'en faut) : on dit que L admet une *décomposition de Zariski algébrique* s'il existe un entier m_0 tel que $m_0 L \simeq \mathcal{O}(E + D)$ où E est un diviseur effectif et D un diviseur nef, en sorte que $H^0(X, kD) \simeq H^0(X, km_0 L)$ pour tout $k \geq 0$. Si $\mathcal{O}(D)$ est semi-ample, il y a alors une métrique lisse à courbure semi-positive sur $\mathcal{O}(D)$, et on en déduit une métrique singulière h sur L de courbure $\frac{1}{m_0}(\Theta(\mathcal{O}(D)) + [E])$, dont les pôles sont constitués du $\mathbb{Q}$-diviseur effectif $\frac{1}{m_0}E$. Pour cette métrique, on a évidemment $\mathcal{I}(h^{km_0}) = \mathcal{O}(-kE)$, de sorte que 4.1.4 est réalisé au moins si m est multiple de m_0.

4.2. Une propriété uniforme de génération globale

Il s'agit d'un résultat qui, d'une certaine manière, donne une "borne inférieure uniforme" pour le produit tensoriel $L \otimes \mathcal{I}(h)$, lorsque h est une métrique singulière quelconque à courbure ≥ 0 sur L.

4.2.1. Proposition. *Soit E un fibré en droites ample au-dessus d'une variété complexe compacte X de dimension n, tel que pour tout point x_0 de X il existe un nombre fini d'éléments de $H^0(X,E)$ qui s'annulent chacun à l'ordre $n+1$ au moins en x_0, et sans zéros communs sur $X \smallsetminus \{x_0\}$. Alors pour tout fibré en droites L sur X ayant une métrique hermitienne singulière localement de la forme $e^{-\xi}$ avec ξ plurisousharmonique d'idéal multiplicateur associé $\mathcal{I}(\xi)$, l'espace des sections globales $H^0(X, \mathcal{O}(K_X+E+L)\otimes\mathcal{I}(\xi))$ engendre le faisceau $\mathcal{O}(K_X+E+L)\otimes\mathcal{I}(\xi)$ en tout point de X.*

Preuve. Nous ne donnerons pas beaucoup de détails. Il y a deux ingrédients clés. D'une part, pour tout germe f de $\mathcal{O}(K_X+E+L)\otimes\mathcal{I}(\xi)$, le théorème de Hörmander-Bombieri permet de trouver une section globale σ pour laquelle

$$\int_{V(x_0)} |f(x)-\sigma(x)|^2 e^{-2\xi} |x-x_0|^{-2(n+1-\varepsilon)} dV < \infty \tag{4.2.2}$$

au voisinage du point donné x_0. Le second ingrédient-clé est le théorème de division L^2 de Skoda [Sko72b, Th. 1, p. 555-556]. Celui-ci permet en effet de déduire de (4.2.2) que $f-\sigma \in \mathcal{O}(K_X+E+L)\otimes\mathcal{I}(\xi)\otimes\mathfrak{m}_{X,x_0}$. On conclut alors par le lemme de Nakayama.

4.3. Le théorème d'extension L^2 de Ohsawa-Takegoshi-Manivel.

Il s'agit d'un théorème de prolongement de sections L^2 de fibrés holomorphes à partir d'une sous-variété de la variété ambiante. Ce théorème fondamental a été d'abord démontré par Ohsawa-Takegoshi [OT87, Ohs88], puis raffiné par L. Manivel [Man93] (au total, la preuve du théorème d'invariance des plurigenres utilise donc 3 types essentiellement différents de théorèmes d'existence L^2 !). Nous énoncerons ici seulement le cas très particulier qui nous intéresse.

4.3.1. Théorème. *Soit $\gamma : X \to \Delta$ une famille projective de variétés complexes compactes paramétrées par le disque ouvert unité $\Delta \subset \mathbb{C}$. Soit $X_0 = \gamma^{-1}(0)$, $n = \dim_{\mathbb{C}} X_0$, et soit L un fibré holomorphe en droites muni d'une métrique hermitienne localement représentée par $e^{-\chi}$, telle que $i\partial\bar\partial\chi \geq \omega$ dans le sens de courants, pour une certaine $(1,1)$-forme positive ω sur X. Soit $0<r<1$ et $\Delta_r = \{t\in\Delta;\ |t|<r\}$. Alors il existe une constante positive A_r ayant la propriété suivante. Pour tout n-forme holomorphe f sur X_0 à valeurs dans L telle que*

$$\int_{X_0} |f|^2 e^{-\chi} dV < \infty,$$

il existe une $(n+1)$-forme holomorphe $\tilde{f}$ sur $\gamma^{-1}(\Delta_r)$ à valeurs dans L, telle que $\tilde{f}|_{X_0} = f \wedge \gamma^*(dt)$ en tout point de X_0 et

$$\int_X |\tilde{f}|^2 e^{-\chi} dV \leq A_r \int_{X_0} |f|^2 e^{-\chi} dV.$$

4.3.2. Remarque. On notera qu'aucune métrique sur le fibré tangent de X_0 ou de X n'est nécessaire pour définir l'intégrale du carré de la norme des formes holomorphes f et $\tilde{f}$ sur X_0 et X ; il suffit en effet d'intégrer les formes volume $i^{n^2} f \wedge \overline{f}$ et $i^{(n+1)^2} \tilde{f} \wedge \overline{\tilde{f}}$.

4.4. Construction de métriques hermitiennes à courbure positive sur K_{X_0}.

On suppose désormais que $X \to \Delta$ est une famille de variétés projectives de type général (i.e. toutes les fibres X_t sont de type général). Un point technique est que la métrique hermitienne sur K_X doit être choisie en sorte que son courant de courbure domine une $(1,1)$-forme définie positive de classe C^∞, de façon à pouvoir appliquer les théorèmes de division et d'extension L^2 de Skoda et Ohsawa-Takegoshi-Manivel. Pour cela, on utilise une variation de la technique de Kodaira qui consiste à écrire un multiple suffisamment grand d'un fibré en droites gros comme somme d'un diviseur effectif et d'un fibré en droites ample.

4.4.1. Lemme. *Il existe un entier positif a tel que $aK_X = D + F$, où D est un diviseur effectif sur X ne contenant pas X_0, et où F est un fibré en droites positif sur X.*

Preuve. Soit F un fibré en droites positif sur X, et soit r_a le rang générique de $H^0(X_t, aK_{X_t} - F)$, qui est atteint pour $t \in \Delta \setminus S_a$ dans le complémentaire d'un ensemble convenable localement fini $S_a \subset \Delta$. Fixons $t_1 \in \Delta \setminus \bigcup S_a$. Puisque X_{t_1} est de type général, nous savons que $h^0(X_{t_1}, aK_{X_{t_1}}) \geq c\, a^n$, donc $h^0(X_{t_1}, aK_{X_{t_1}} - F) \geq c' a^n$ pour $c, c' > 0$ convenables et a assez grand. D'après le choix de t_1, toute section non nulle de $H^0(X_{t_1}, aK_{X_{t_1}} - F)$ s'étend en une section s de $H^0(X, aK_X - F)$, donc $aK_X = D + F$ où D est le diviseur de s. Si nécessaire, on peut éliminer la composante X_0 dans la décomposition de D en divisant s par une puissance convenable de t. □

L'étape suivante est de construire des métriques hermitiennes sur K_{X_0} et K_X, respectivement, et de comparer leurs faisceaux d'idéaux multiplicateurs. D'après la section 4.1, il existe une métrique à singularités minimales sur X_0 (unique à équivalence de singularité près). On notera φ_0 le poids de cette métrique. De même, il existe une métrique à singularités minimales sur tout voisinage relativement compact

de X_0 dans X, et quitte à rétrécir la base Δ on peut supposer que cette métrique existe sur l'espace total X tout entier. On notera φ *la restriction à* X_0 du poids de cette métrique. Par définition, φ est au moins aussi singulier que φ_0 sur X_0, et après ajout éventuel d'une constante on peut supposer que $\varphi \leq \varphi_0$.

Par ailleurs, d'après le lemme 4.4.1, on peut choisir un entier $a \geq 2$ tel que $aK_X = D + F$, où D est un diviseur effectif sur X ne contenant pas X_0, et F un fibré en droites positif sur X. Quitte à remplacer a par un multiple assez grand, on peut faire les hypothèses supplémentaires suivantes:

(4.4.2) *pour tout* $x_0 \in X_0$, *il existe un nombre fini d'éléments de* $H^0(X, F - 2K_X)_{|X_0}$ *dont le lieu des zéros communs est réduit au singleton* $\{x_0\}$, *et s'annulant à un ordre au moins égal à* $n+1$ *en* x_0.

(4.4.3) *une base des sections de* $H^0(X,F)_{|X_0}$ *fournit un plongement de* X_0 *sur une sous-variété d'un espace projectif complexe.*

Soit s_D la section canonique du fibré en droites $\mathcal{O}(D)$, de sorte que le diviseur de s_D est D. Soient $u_1, \ldots, u_N \in H^0(X,F)$ des sections telles que

$$u_{1|X_0}, \ldots, u_{N|X_0}$$

forment une base de $H^0(X,F)_{|X_0}$. Puisque $s_D u_j \in H^0(X, aK_X)$ $(1 \leq j \leq N)$, on obtient une métrique hermitienne

$$e^{-\psi} = \left(\frac{1}{|s_D|^2 \sum |u_j|^2} \right)^{\frac{1}{a}}$$

sur le fibré en droites K_X. De plus

$$\psi = \frac{1}{a} \left(\log |s_D|^2 + \chi \right)$$

où $\chi = \log(\sum |u_j|^2)$ définit une métrique hermitienne de classe C^∞ à courbure positive sur F. Par conséquent, $i\partial\overline{\partial}\psi$ est un courant défini positif. En outre, les singularités de ψ sur X sont au moins aussi grandes que celles du poids φ qui a les singularités minimales sur X. En ajustant de nouveau les constantes si nécessaire, on peut énoncer le

4.4.4. Lemme. *La métrique hermitienne* $e^{-\psi}$ *est à courbure définie positive et* $\psi \leq \varphi \leq \varphi_0$ *sur* X_0.

L'argument crucial est un résultat de comparaison des faisceaux d'idéaux multiplicateurs sur X_0 définis par $\ell\varphi_0$ et $\ell\varphi$, respectivement, lorsque ℓ est grand. Dans tout

le reste de cet exposé, la notation $\mathcal{I}((\ell+a-\varepsilon)\varphi+\varepsilon\psi)$ désignera un faisceau d'idéaux sur X_0 (et non pas un faisceau d'idéaux sur X, même si le poids est éventuellement défini sur X tout entier).

4.4.5. Proposition. *Choisissons $0<\varepsilon<1$ assez petit pour que $e^{-\varepsilon\psi}$ soit localement intégrable sur X (peut-être après avoir un peu rétréci le disque Δ) et pour que $e^{-\varepsilon\psi}_{|X_0}$ soit localement intégrable sur X_0. Alors*

$$\mathcal{I}((\ell-\varepsilon)\varphi_0+(a+\varepsilon)\psi)\subset\mathcal{I}((\ell-1+a-\varepsilon)\varphi+\varepsilon\psi)$$

pour tout entier $\ell\geq 1$.

Preuve. On raisonne par récurrence sur ℓ. Pour $\ell=1$, le Lemme 4.4.4 implique

$$(1-\varepsilon)\varphi_0+(a+\varepsilon)\psi\leq(1-\varepsilon)\varphi_0+a\varphi+\varepsilon\psi\leq C+(a-\varepsilon)\varphi+\varepsilon\psi$$

puisque φ_0 est localement majorée par une certaine constante C. Nous obtenons donc

$$\mathcal{I}((1-\varepsilon)\varphi_0+(a+\varepsilon)\psi)\subset\mathcal{I}((a-\varepsilon)\varphi+\varepsilon\psi),$$

comme désiré.

Maintenant, supposons que l'inclusion ait été démontrée pour l'entier ℓ. Prenons un germe de fonction f quelconque dans le faisceau d'idéaux $\mathcal{I}((\ell+1-\varepsilon)\varphi_0+(a+\varepsilon)\psi)$, défini sur un petit voisinage U d'un point $P\in X_0$. Fixons un repère holomorphe local e de $(\ell+a)K_{X_0}$ sur U. Alors $s=fe$ est une section de

$$\mathcal{O}((\ell+a)K_{X_0})\otimes\mathcal{I}((\ell+1-\varepsilon)\varphi_0+(a+\varepsilon)\psi)$$

sur U. Observons qu'étant donné une fonction plurisousharmonique arbitraire ξ, on a

$$\mathcal{I}(\xi+\log|s_D|^2)=\mathcal{I}(\xi)\otimes\mathcal{O}(-D).$$

En écrivant $aK_{X_0}=(D+F)_{|X_0}$ grâce au Lemme 4.4.1 et $a\psi=\log|s_D|^2+\chi$ par définition de ψ, on peut réinterpréter s comme étant une section de

$$\begin{aligned}\mathcal{O}(\ell K_{X_0}+(D+F)_{|X_0})\otimes\mathcal{I}((\ell+1-\varepsilon)\varphi_0+\varepsilon\psi+\log|s_D|^2+\chi)&\\ =\mathcal{O}((\ell+2)K_{X_0}+E_{|X_0})\otimes\mathcal{I}((\ell+1-\varepsilon)\varphi_0+\varepsilon\psi)&\end{aligned}$$

où $E=F-2K_X$ (comme χ est C^∞, χ ne change pas le faisceau multiplicateur). Observons que $(\ell+1-\varepsilon)\varphi_0+\varepsilon\psi$ définit une métrique hermitienne à courbure positive

sur $(\ell+1)K_{X_0}$. D'après l'hypothèse (4.4.2) ci-dessus et la Proposition 4.2.1 appliquée avec $Y = X_0$ et $L = (\ell+1)K_{X_0}$, on conclut que

$$\mathcal{O}((\ell+2)K_{X_0} + E_{|X_0}) \otimes \mathcal{I}((\ell+1-\varepsilon)\varphi_0 + \varepsilon\psi)$$

est engendré par ses sections globales sur X_0. On peut donc sans perte de généralité se restreindre au cas des germes f tels que fe coïncide sur U avec une section globale

$$s \in H^0\left(X_0, \mathcal{O}((\ell+2)K_{X_0} + E_{|X_0}) \otimes \mathcal{I}((\ell+1-\varepsilon)\varphi_0 + \varepsilon\psi)\right).$$

En renversant le sens des calculs ci-dessus, on trouve

$$\begin{aligned} H^0\left(X_0, \mathcal{O}((\ell+2)K_{X_0}\right. &\left. + E_{|X_0}) \otimes \mathcal{I}((\ell+1-\varepsilon)\varphi_0 + \varepsilon\psi)\right) \\ &= H^0\left(X_0, \mathcal{O}((\ell+a)K_{X_0}) \otimes \mathcal{I}((\ell+1-\varepsilon)\varphi_0 + (a+\varepsilon)\psi)\right) \\ &\subset H^0\left(X_0, \mathcal{O}((\ell+a)K_{X_0}) \otimes \mathcal{I}((\ell-\varepsilon)\varphi_0 + (a+\varepsilon)\psi)\right) \\ &\subset H^0\left(X_0, \mathcal{O}((\ell+a)K_{X_0}) \otimes \mathcal{I}((\ell+a-1-\varepsilon)\varphi + \varepsilon\psi)\right), \end{aligned}$$

[la première inclusion est obtenue en oubliant le terme φ_0 dans le poids, et la seconde est une conséquence de l'hypothèse de récurrence pour ℓ]. Maintenant, le poids

$$(\ell+a-1-\varepsilon)\varphi + \varepsilon\psi$$

définit une métrique hermitienne à courbure définie positive sur $L = (\ell+a-1)K_X$, et la Proposition 4.3.1 implique que s s'étend en une section globale $\tilde{s} \in H^0(X, (\ell+a)K_X)$ (peut-être après avoir rétréci un peu Δ). La définition de φ implique $|\tilde{s}|^2 \leq C\, e^{(\ell+a)\varphi}$, donc

$$|\tilde{s}|^2 e^{-(\ell+a)\varphi - \varepsilon\psi} \leq C\, e^{-\varepsilon\psi}$$

est intégrable sur X. De là, on conclut que

$$fe = s = \tilde{s}_{|X_0} \in \mathcal{O}((\ell+a)K_{X_0}) \otimes \mathcal{I}((\ell+a)\varphi + \varepsilon\psi),$$

donc $f \in \mathcal{I}((\ell+a)\varphi + \varepsilon\psi)$. L'étape $\ell+1$ de la récurrence est démontrée.

4.5. Preuve du théorème d'invariance des plurigenres.

Fixons un entier $m > 0$ et une section $s \in H^0(X_0, mK_{X_0})$ quelconque. Par définition de φ_0, on a $|s|^2 \leq Ce^{m\varphi_0}$ sur X_0. Si s_D est la section canonique de $\mathcal{O}(-D)$ de diviseur D, on en déduit que $s^\ell s_D$ est localement L^2 par rapport au poids $e^{-\ell m\varphi_0 - (a+\varepsilon)\psi}$, car $a\psi$ possède la même singularité que $\log|s_D|^2$ et $e^{-\varepsilon\psi}$ est supposée

localement intégrable sur X_0. Les fonctions trivialisant $s^\ell s_D$ sont localement dans le faisceau d'idéaux $\mathcal{I}((\ell m-\varepsilon)\varphi_0+(a+\varepsilon)\psi)$. D'après la Proposition 4.4.5, elles appartiennent à $\mathcal{I}((\ell m-1-\varepsilon)\varphi+\varepsilon\psi)$, i.e.

$$\int_U |s^\ell s_D|^2 e^{-(\ell m-1-\varepsilon)\varphi-\varepsilon\psi} < +\infty$$

sur tout ouvert U suffisamment petit tel que $K_{X_0}|U$ et $\mathcal{O}(-D)|U$ soient triviaux. D'une manière équivalente, on peut écrire

$$\int_U |s|^{2\ell} e^{-(\ell m-1-\varepsilon)\varphi+(a-\varepsilon)\psi} < +\infty.$$

Prenons ℓ assez grand pour que $a/(\ell-1)\le\varepsilon$, de sorte que $e^{\varphi-\frac{a}{\ell-1}\psi}\le Ce^{-\varepsilon\psi}$ soit intégrable sur U. D'après l'inégalité de Hölder d'exposants conjugués ℓ, $\ell'=\ell/(\ell-1)$, on trouve

$$+\infty > \left(\int_U |s|^{2\ell} e^{-(\ell m-1-\varepsilon)\varphi+(a-\varepsilon)\psi}\right)^{1/\ell} \left(\int_U e^{\varphi-\frac{a}{\ell-1}\psi}\right)^{(\ell-1)/\ell}$$

$$\ge \int_U |s|^2 e^{-(m-\frac{1}{\ell}-\frac{\varepsilon}{\ell})\varphi+(\frac{a}{\ell}-\frac{\varepsilon}{\ell})\psi} e^{(1-\frac{1}{\ell})\varphi-\frac{a}{\ell}\psi} = \int_U |s|^2 e^{-(m-1-\delta)\varphi-\delta\psi}$$

avec $\delta=\varepsilon/\ell$. On peut considérer s comme une section de $K_{X_0}+L_{|X_0}$ avec $L=(m-1)K_X$. Le poids $(m-1-\delta)\varphi-\delta\psi$ définit une métrique hermitienne sur L à courbure définie positive, et s est globalement L^2 par rapport à cette métrique. D'après le théorème d'extension de Ohsawa-Takegoshi-Manivel, on peut par conséquent étendre s en une section $\tilde{s}\in H^0(X,K_X+L)=H^0(X,mK_X)$, éventuellement après avoir tronqué Δ. Ceci termine la preuve du théorème.

BIBLIOGRAPHIE

[AN54] Y. Akizuki and S. Nakano - *Note on Kodaira-Spencer's proof of Lefschetz theorems*, Proc. Jap. Acad. **30** (1954), 266–272.

[AS95] U. Angehrn and Y.T. Siu - *Effective freeness and point separation for adjoint bundles*, Invent. Math. **122** (1995), 291–308.

[AV65] A. Andreotti and E. Vesentini - *Carleman estimates for the Laplace-Beltrami equation in complex manifolds*, Publ. Math. I.H.E.S. **25** (1965) 81–130.

[BeS93] M. Beltrametti and A.J. Sommese - *On k-jet ampleness*, Complex Analysis and Geometry, Univ. Series in Math., edited by V. Ancona and A. Silva, Plenum Press, New-York, (1993), 355–376.

[BFS89] M. Beltrametti, P. Francia and A.J. Sommese - *On Reider's method and higher order embeddings*, Duke Math. J. **58** (1989), 425–439.

[Boc48] S. Bochner - *Curvature and Betti numbers (I) and (II)*, Ann. of Math. **49** (1948), 379–390; **50** (1949), 77–93.

[Bom70] E. Bombieri - *Algebraic values of meromorphic maps*, Invent. Math. **10** (1970), 267–287 and *Addendum*, Invent. Math. **11** (1970), 163–166.

[Bon93] L. Bonavero - *Inégalités de Morse holomorphes singulières*, C. R. Acad. Sci., Paris, Sér. I 317, n° 12, 1163-1166 (1993)

[Dem82] J.-P. Demailly - *Estimations L^2 pour l'opérateur $\overline{\partial}$ d'un fibré vectoriel holomorphe semi-positif au-dessus d'une variété kählérienne complète*, Ann. Sci. Ec. Norm. Sup. **15** (1982), 457–511.

[Dem85] J.-P. Demailly - *Champs magnétiques et inégalités de Morse pour la d''-cohomologie*, Ann. Inst. Fourier (Grenoble) **35** (1985), 189–229.

[Dem89] J.-P. Demailly - *Transcendental proof of a generalized Kawamata-Viehweg vanishing theorem*, C. R. Acad. Sci. Paris Sér. I Math. **309** (1989), 123–126 and:
Proceedings of the Conference "Geometrical and algebraical aspects in several complex variables" held at Cetraro, Univ. della Calabria, June (1989).

[Dem90] J.-P. Demailly - *Singular hermitian metrics on positive line bundles*, Proc. Conf. Complex algebraic varieties (Bayreuth, April 2–6, 1990), edited by K. Hulek, T. Peternell, M. Schneider, F. Schreyer, Lecture Notes in Math., Vol. 1507, Springer-Verlag, Berlin, (1992).

[Dem93] J.-P. Demailly - *A numerical criterion for very ample line bundles*, J. Differential Geom. **37** (1993), 323–374.

[Dem96] J.-P. Demailly - *Effective bounds for very ample line bundles*, Invent. Math. **124** (1996), 243–261.

[Dem98] J.-P. Demailly - *Pseudoconvex-concave duality and regularization of currents*, Prépublication Institut Fourier, Septembre 1998, 40 p.

[DK99] J.-P. Demailly, J. Kollár - *Semi-continuity of complex singularity exponents and Kähler-Einstein metrics on Fano manifolds*, prépublication Institut Fourier 1999.

[DPS94] J.-P. Demailly, Th. Peternell, M. Schneider - *Compact complex manifolds with numerically effective tangent bundles*, J. Algebraic Geometry **3** (1994)

295–345.

[DPS98] J.-P. Demailly, Th. Peternell, M. Schneider - *Pseudo-effective line bundles on projective varieties*, manuscript in preparation (1998).

[Ein94] L. Ein - *Note on higher dimensional adjoint linear systems*, preprint 1994, personal communication to the author.

[EKL95] L. Ein, O. Küchle and R. Lazarsfeld - *Local positivity of ample line bundles*, J. Differ. Geom. 42, (1995), 193-219.

[EL92] L. Ein and R. Lazarsfeld - *Seshadri constants on smooth surfaces*, Journées de Géométrie Algébrique d'Orsay, Juillet 1992, Astérisque **282** (1993), 177–186.

[EL93] L. Ein and R. Lazarsfeld - *Global generation of pluricanonical and adjoint linear series on smooth projective threefolds*, Jour. of Am. Math. Soc. **6** (1993), 875–903.

[EL95] L. Ein and R. Lazarsfeld - *Global generation of linear series on terminal threefolds*, Intern. J. Math. **6** (1995) 1–18.

[EV86] H. Esnault and E. Viehweg - *Logarithmic De Rham complexes and vanishing theorems*, Invent. Math. **86** (1986) 161–194.

[EV92] H. Esnault and E. Viehweg - *Lectures on vanishing theorems*, DMV Seminar, Band **20**, Birkhäuser Verlag (1992).

[FdB95] G. Fernández del Busto - *Bogomolov instability and Kawamata-Viehweg vanishing*, J. Algebr. Geom. **4** (1995) 693–700.

[FdB96] G. Fernández del Busto - *A Matsusaka-type theorem on surfaces*, J. Algebr. Geom. **5** (1996), 513–520.

[Fuj83] T. Fujita - *Semipositive line bundles*, J. Fac. Sci. Univ. of Tokyo **30** (1983), 353–378.

[Fuj87] T. Fujita - *On polarized manifolds whose adjoint bundles are not semipositive*, Algebraic Geometry, Sendai, 1985, Adv. Stud. in Pure Math., Vol. 10, North Holland, T. Oda (ed.) (1987), 167–178.

[Fuj88] T. Fujita - *Problem list*, Conference held at the Taniguchi Foundation, Katata, Japan, August 1988.

[Fuj93] T. Fujita - *Remarks on Ein-Lazarsfeld criterion of spannedness of adjoint bundles of polarized threefolds*, preprint 1993.

[Gra60] H. Grauert - *Ein Theorem der analytischen Garbentheorie und die Modulräume komplexer Strukturen*, Publ. Math. I.H.E.S. **5** (1960), 233–292.

[Gri69] P.A. Griffiths - *Hermitian differential geometry, Chern classes and positive vector bundles*, Global Analysis, papers in honor of K. Kodaira, Princeton

Univ. Press, Princeton, 1969, 181–251.

[GH78] P.A. Griffiths, J. Harris - *Principles of algebraic geometry*, Wiley, New York (1978).

[Hel97] S. Helmke - *On Fujita's conjecture*, Duke Math. J. **88** (1997), 201–231.

[Hel98] S. Helmke - *Global generation of adjoint linear systems*, Communication au Congrès d'Oberwolfach "Komplexe Analysis", Septembre 1998.

[Hir64] H. Hironaka - *Resolution of singularities of an algebraic variety over a field of characteristic zero*, Ann. of Math. **79** (1964) 109–326.

[Hör65] L. Hörmander - *L^2 estimates and existence theorems for the $\overline{\partial}$ operator*, Acta Math. **113** (1965), 89–152.

[Hör66] L. Hörmander - *An introduction to Complex Analysis in several variables*, 1966, 3rd edition, North-Holland Math. Libr., vol.7, Amsterdam, London (1990).

[Kaw82] Y. Kawamata - *A generalization of Kodaira-Ramanujam's vanishing theorem*, Math. Ann. **261** (1982), 43–46.

[Kaw84] Y. Kawamata - *The cone of curves of algebraic varieties*, Ann. of Math. **119** (1984) 603–633.

[Kaw97a] Y. Kawamata - *On Fujita's freeness conjecture for 3-folds and 4-folds*, Math. Ann. **308** (1997), 491–505.

[Kaw97b] Y. Kawamata - *Deformation of canonical singularities*, alg-geom/9712018 (1997), to appear in J. Amer. Math. Soc.

[Kaw98] Y. Kawamata - *On the extension problem of pluricanonical forms*, preprint Univ. of Tokyo at Komaba, September 1998.

[KMM87] Y. Kawamata, K. Matsuda and K. Matsuki - *Introduction to the minimal problem* Algebraic Geometry, Sendai, 1985, Adv. Stud. in Pure Math., Vol. 10, T. Oda (ed.), North Holland, Amsterdam (1987), 283–360.

[Kod53] K. Kodaira - *On a differential geometric method in the theory of analytic stacks*, Proc. Nat. Acad. Sci. USA **39** (1953), 1268–1273

[Kod54] K. Kodaira - *On Kähler varieties of restricted type*, Ann. of Math. **60** (1954), 28–48.

[Kol92] J. Kollár - *Effective basepoint freeness*, Math. Ann. **296** (1993), 595–605.

[Kol95] J. Kollár - *Singularity of pairs*, Preprint University of Utah, November 1995.

[KoM83] J. Kollár and T. Matsusaka - *Riemann-Roch type inequalities*, Amer. J. of Math. **105** (1983), 229–252.

[KoMM92] J. Kollár, Y. Miyaoka and S. Mori - *Rational connectedness and boundedness of Fano manifolds*, J. Differential Geom. **36** (1992), 765–779.

[Laz93] R. Lazarsfeld, - *Lectures on linear series, with the assistance of G. Fernández del Busto*, Lectures of a summer program on Complex Algebraic Geometry, Park City, 1993, Amer. Math. Soc., IAS/Park City Math. Ser. **3** (1997), 163–219.

[Lel57] P. Lelong - *Intégration sur un ensemble analytique complexe*, Bull. Soc. Math. France **85** (1957), 239–262.

[Lel69] P. Lelong - *Plurisubharmonic functions and positive differential forms*, Gordon and Breach, New-York, and Dunod, Paris (1969).

[Man93] L. Manivel - *Un théorème de prolongement L^2 de sections holomorphes d'un fibré vectoriel*, Math. Zeitschrift **212** (1993) 107–122.

[Mat72] T. Matsusaka - *Polarized varieties with a given Hilbert polynomial*, Amer. J. of Math. **94** (1972), 1027–1077.

[Mor82] S. Mori - *Threefolds whose canonical bundles are not numerically effective*, Ann. of Math. **116** (1982), 133–176.

[Nad89] A.M. Nadel - *Multiplier ideal sheaves and Kähler-Einstein metrics of positive scalar curvature*, Proc. Nat. Acad. Sci. U.S.A. **86** (1989), 7299–7300 and Annals of Math., **132** (1990), 549–596.

[Nak55] S. Nakano - *On complex analytic vector bundles*, J. Math. Soc. Japan **7** (1955), 1–12.

[Ohs88] T. Ohsawa - *On the extension of L^2 holomorphic functions, II*, Publ. RIMS, Kyoto Univ. **24** (1988), 265–275.

[OT87] T. Ohsawa and K. Takegoshi - *On the extension of L^2 holomorphic functions*, Math. Zeitschrift **195** (1987), 197–204.

[Rei88] I. Reider - *Vector bundles of rank 2 and linear systems on algebraic surfaces*, Ann. of Math. **127** (1988), 309–33.1

[Rh55] G. De Rham - *Variétés différentiables*, Hermann, Paris (1955).

[Sak88] F. Sakai - *Reider-Serrano's method on normal surfaces*, Proc. Intern. Conf. on Algebraic Geometry (L'Aquila, June 1988), Lecture Notes in Math., Vol. 1417, Springer-Verlag, Berlin, 1990, 301–319.

[Sch74] M. Schneider - *Ein einfacher Beweis des Verschwindungssatzes für positive holomorphe Vektorraumbündel*, Manuscripta Math. **11** (1974), 95–101.

[ShSo85] B. Shiffman, A.J. Sommese - *Vanishing theorems on complex manifolds*, Progress in Math. no **56**, Birkhäuser (1985).

[Sho85] V. Shukorov - *The non-vanishing theorem*, Math. U.S.S.R. Izv. **19** (1985), 591–607.

[Siu74] Y.T. Siu - *Analyticity of sets associated to Lelong numbers and the extension*

of closed positive currents, Invent. Math. **27** (1974), 53–156.

[Siu93] Y.T. Siu - *An effective Matsusaka big theorem*, Ann. Inst. Fourier **43** (1993), 1387–1405.

[Siu95] Y.T. Siu - *Very ampleness criterion of double adjoint of ample line bundles*, Princeton Conf. in honor of Robert C. Gunning and Joseph J. Kohn, Princeton University, NJ, USA, Mar. 16-20, 1992; Ann. Math. Stud. **137** (1995), 291–318.

[Siu96] Y.T. Siu - *Effective Very Ampleness*, Invent. Math. **124** (1996), 563–571.

[Siu97] Y.T. Siu - *Invariance of Plurigenera*, manuscript October 1997, to appear in Inventiones Math.

[Sko72a] H. Skoda - *Sous-ensembles analytiques d'ordre fini ou infini dans* $\mathbb{C}^n$, Bull. Soc. Math. France **100** (1972), 353–408.

[Sko72b] H. Skoda - *Applications des techniques* L^2 *à la théorie des idéaux d'une algèbre de fonctions holomorphes avec poids*, Ann. Scient. Ec. Norm. Sup. 4e Série **5** (1972), 545–579.

[Sko75] H. Skoda - *Estimations* L^2 *pour l'opérateur* $\overline{\partial}$ *et applications arithmétiques*, Séminaire P. Lelong (Analyse), année 1975/76, Lecture Notes in Math., Vol. 538, Springer-Verlag, Berlin (1977), 314–323.

[Tsu96] H. Tsuji - *Global generation of adjoint bundles*, Nagoya Math. J. **142** (1996), 5–16.

[Vie82] E. Viehweg - *Vanishing theorems*, J. Reine Angew. Math. **335** (1982), 1–8.

Jean–Pierre Demailly

Université de Grenoble I
Institut Fourier, UMR 5582 du CNRS
B.P. 74, route des Maths
F–38402 Saint-Martin-d'Hères Cedex

E–mail : Jean-Pierre.Demailly@ujf-grenoble.fr

Astérisque

TOMASZ KOMOROWSKI

Brownian motion in a Poisson obstacle field

Astérisque, tome 266 (2000), Séminaire Bourbaki, exp. n° 853, p. 91-111
<http://www.numdam.org/item?id=SB_1998-1999__41__91_0>

© Société mathématique de France, 2000, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (http://smf4.emath.fr/Publications/Asterisque/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM
Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Séminaire BOURBAKI
51ème année, 1998-99, n° 853

Novembre 1998

BROWNIAN MOTION IN A POISSON OBSTACLE FIELD

by **Tomasz KOMOROWSKI**

1. INTRODUCTION

The problem of a diffusion performed in a medium with randomly distributed static traps appeared in Physics literature some 80 years ago (cf. e.g. Smoluchowski (1916)). A common illustration is a motion of a Brownian particle among soft or hard core obstacles of finite radius with centers Poisson distributed in a d dimensional space, cf. e.g. Grassberger et al. (1982).

It can be shown, cf. Section 2, that on the average the behavior of the survival probability S_t of the particle in the field up to time t, the so called *annealed asymptotic*, is given by

$$S_t = \exp\left\{-\tilde{c}(d,\nu)t^{\frac{d}{d+2}}(1+o(1))\right\},\ t \gg 1.$$

Here ν is the intensity of the Poisson cloud and $\tilde{c}(d,\nu)$ is a constant given by (8). This result has been obtained non-rigorously in Kac and Luttinger (1973-74). The decay rate of the survival probability is much slower than exponential. The latter could be expected in light of the result of Kesten, Spitzer, Whitman, cf. the remark after Theorem 1.

The problem of determining the annealed asymptotic is related, as we shall explain in Section 5 below, to the question of finding a rigorous proof of the so called Lifshitz tail effect for the Integrated Density of States (IDS) function of the Schrödinger operator with a random Poisson potential. This question appeared in Physics literature in the context of quantum theory of a condensed state in the work of Lifshitz, cf. Lifshitz (1965).

The first rigorous argument giving the correct asymptotic of S_t via large deviation theory of Wiener sausage asymptotic has been presented by Donsker and Varadhan (1975).

This result has been obtained again via a different technique – the so called *Method of Enlargement of Obstacles* (MEO) in Sznitman (1990). The method later became a fundamental tool in understanding a rich collection of phenomena associated with Brownian motion among Poissonian traps such as: the asymptotic of the survival probability $S_{t,\omega}$

in a typical realization of the medium ω - the so called *quenched asymptotic*, the *Pinning Effect* and *Confinement Property* for the trapped Brownian Motion, cf. Sznitman (1998) and (1991), Povel (1998), the Lifshitz tails for the IDS function of the random magnetic Schrödinger operator with Poissonian random potential, cf. Erdösz (1998) to name just a few. Other results go in the direction of determining the behavior of the Brownian particle performing a long crossing between two distant points of the medium and consequently lead to the introduction of the so called Lyapunov exponents, cf. Sznitman (1994), (1995a,b) and also the Shape Theorem of Section 3.1 for the quenched case. The Lyapunov exponents turn out to be very convenient tools in deriving large deviation results for the quenched and annealed measures associated with the surviving Brownian process, cf. Sznitman (1995), (1995a,b).

In the context of Lyapunov exponent one can introduce also the notion of Crossing Brownian Motion (CBM), cf. Section 4. CBM measures appear naturally in certain physical models of growing interfaces, cf. Krug-Spohn (1991). One of the basic questions is to understand the relation between the size of transverse fluctuations of the CBM and the fluctuations of a certain natural random distance function defined by (17). This direction has been pursued by Wüthrich in Wüthrich (1998).

Lyapunov exponents have been also applied to the analysis of random walks in a random environment (RWRE) in higher dimensions. The appropriate Shape Theorem has been proven in this context by Zerner (1997). He used the resulting Lyapunov exponents in the derivation of the large deviation theorem for random walks in dimension $d \geq 2$ satisfying the so called *nestling property.*

To describe rigorously the obstacle field we consider $\mathbb{P}$ to be a probability measure given on Ω – the set of locally finite simple pure point measures on $\mathbb{R}^d$

$$\omega = \sum_i \delta_{x_i},\ x_i \in \mathbb{R}^d \tag{1}$$

endowed with the canonical σ-algebra generated by the mappings

$$\omega \in \Omega \to \omega(A) \in \mathbb{Z}_+,\ A \in \mathcal{B}(\mathbb{R}^d) \text{ - the Borelian } \sigma\text{-algebra.}$$

We shall assume that the centers of the obstacles form a cloud of Poisson distributed points with intensity $\nu > 0$, i.e. that

$$\mathbb{E}\exp\left\{-\int f(y)\omega(dy)\right\} = \exp\left\{-\nu\int dy\left(1-e^{-f(y)}\right)\right\}$$

for any Borel measurable function f. Here $\mathbb{E}$ denotes the mathematical expectation of measure $\mathbb{P}$.

Let $Z.$ be the canonical d-dimensional Brownian Motion, P_x the Wiener measure corresponding to a Brownian particle starting at x and E_x its corresponding expectation. We shall assume that the particle interacts with the cloud ω via a potential representing the

absorption rate of the random medium

$$V(x,\omega) := \sum_i W(x - x_i, \omega),$$

where x_i are determined from (1) and W is a nonnegative, measurable function whose support is contained in a ball of radius $a > 0$ centered at 0.

In what follows we shall also discuss the hard obstacle case. We define then the obstacle set as $A := \bigcup_{x_i}(x_i + K)$ for some nonpolar set K (i.e. of positive capacity) of finite diameter a. The particle is instantaneously killed upon entry into the region

$$\Theta := \mathbb{R}^d - A. \tag{2}$$

Informally speaking, in the hard obstacle case, we take $V(\cdot,\omega) = \infty$ on the obstacle set and 0 elsewhere.

The structure of our presentation is as follows. In the first four sections we give a brief overview of some of the results available in the area. Because of the scope of this paper we are unable to provide more than a short sketch of the techniques used in the field. However we try our best to give the reader some, mostly heuristic, explanations of the presented results while avoiding getting too technical in our presentation.

Section 6 presents the main ingredients of MEO presented here after Sznitman (1998), which, as we have already mentioned, is crucial in the theory. Here we try to be more precise than in the previous sections although we maintain our goal of keeping the presentation as simple as possible.

Finally we close the article with a short review of some open problems in the subject.

2. QUENCHED AND ANNEALED MEASURES

The Brownian motion among the obstacles can be described by the path measures

$$Q_{t,\omega} := \frac{1}{S_{t,\omega}} \exp\left\{-\int_0^t V(Z_s,\omega)ds\right\} P_0 \qquad \textit{the quenched measure} \tag{3}$$

with $\omega \in \Omega$ or

$$Q_t := \frac{1}{S_t} \exp\left\{-\int_0^t V(Z_s,\omega)ds\right\} P_0 \otimes \mathbb{P} \qquad \textit{the annealed measure.} \tag{4}$$

Here the normalizing constants are given by

$$S_{t,\omega} := E_0\left[\exp\left\{-\int_0^t V(Z_s,\omega)ds\right\}\right]$$

and

$$S_t := E_0 \otimes \mathbb{E}\left[\exp\left\{-\int_0^t V(Z_s,\omega)ds\right\}\right].$$

Remark. Let us observe that in the hard obstacle case we have

$$S_t = E_0[\exp\{-\nu|W_t^a|\}], \tag{5}$$

where $W_t^a := \bigcup_{0\le s\le t} Z_s + \overline{B}(0,a)$ is the tubular neighborhood of radius a of the Brownian trajectory, the so called "Wiener sausage".

Intuitively speaking the quenched measure describes the law of the particle conditioned on the event that it survives among the traps up to time t for a typical configuration of traps ω. On the other hand the annealed measure arises as a result of averaging over possible realisations of the random medium.

The question of interest is the asymptotic behavior of the path measures (3), (4) when $t \uparrow +\infty$. In the first step we describe the asymptotics of $S_{t,\omega}$ and S_t – the survival probabilities of the Brownian Motion among the obstacles – for large t. The following result holds.

THEOREM 1. — *When $t \uparrow +\infty$ $\mathbb{P}$ a.s. we have*

$$S_{t,\omega} = \exp\left\{-c(d,\nu)\frac{t}{(\log t)^{2/d}}(1+o(1))\right\} \text{ for soft obstacles} \tag{6}$$

and

$$S_t = \exp\left\{-\tilde{c}(d,\nu)t^{\frac{d}{d+2}}(1+o(1))\right\} \text{ for both soft and hard obstacles.} \tag{7}$$

Here

$$c(d,\nu) := \lambda_d\left(\frac{\nu\omega_d}{d}\right)^{2/d}$$

and

$$\tilde{c}(d,\nu) := \inf_{U \text{ open}}\{\nu|U| + \lambda(U)\} = (\nu\omega_d)^{2/d+2}\left(\frac{d+2}{2}\right)\left(\frac{2\lambda_d}{d}\right)^{\frac{d}{d+2}}, \tag{8}$$

with $\lambda(U)$ denoting the principal Dirichlet eigenvalue of the Laplacian $-\frac{1}{2}\Delta$ corresponding to the region U. We have written λ_d and ω_d for the fundamental tone and volume of the unit ball in the d dimensional space respectively.

Remark. At this time it may be worthwhile to mention that motivated by the result of Kesten, Spitzer and Whitman (cf. Spitzer (1964), p. 40) stating that $\frac{|W_t^a|}{t}$ tends to the capacity of $B(0,a)$ (the ball of radius a centered at 0) when $t \uparrow +\infty$ and (5) one could mistakenly believe that S_t decays exponentially, which obviously contradicts (7).

The following argument gives some plausibility to the lower bounds on the asymptotic of the normalizing constants. In Section 6 we shall sketch the proof of the upper bound on S_t using MEO.

Informally speaking (when W is sufficiently smooth this is in fact accurate) we can write, using Feynman-Kac formula, that

$$S_{t,\omega} = u_\omega(t,0) \qquad \text{and} \qquad S_t = \mathbb{E}[u_\omega(t,0)],$$

where

$$u_\omega(t,x) := E_x \exp\left\{-\int_0^t V(Z_s,\omega)ds\right\}$$

satisfies the Cauchy problem

$$\partial_t u_\omega(t,x) = -H_\omega u_\omega(t,x), \tag{9}$$

$$u_\omega(0,x) \equiv 1$$

with

$$H_\omega := -\frac{1}{2}\Delta + V(x,\omega). \tag{10}$$

Using (9) we obtain that

$$S_{t,\omega} \approx \sum_i e^{-\lambda_i t} < \varphi_i, 1 > \varphi_i(0) \tag{11}$$

where λ_i, $i = 1, 2, \cdots$ denotes the complete set of eigenvalues of the self-adjoint operator H_ω and φ_i are the corresponding eigenfunctions. Consider now a sufficiently large box $\Lambda = (-l, l)^d$. With probability $\mathbb{P}$ close to one Λ contains a ball $B_l := B(x_l, R_l)$ of radius $R_l \sim R_0(\log l)^{1/d}$ on which $V(\cdot,\omega) \equiv 0$, $R_0 = \left(\frac{d}{\nu\omega_d}\right)^{1/d}$. In that case we have an easy upper bound $\lambda_1 \leq \lambda(B_l)$ – the Dirichlet principal eigenvalue of the Laplacian operator corresponding to the ball B_l. On the other hand we should expect then φ_1 to be localized in B_l and decay exponentially at a rate α away from x_l which is a consequence of e.g. formula (2.26) p. 276 of Sznitman (1998). Thus we can suppose that $\varphi_1(0) \sim e^{-\alpha|x_l|}$. Discarding the influence of higher eigenvalues we see that asymptotically $S_{t,\omega}$ should be greater than or equal to $\exp\{-(!!\alpha|x_l| + \lambda(B_l)t)\}$. Choosing now $l = o(\frac{t}{(\log t)^{2/d}})$ but such that $\log l \sim \log t$, we get

$$S_{t,\omega} \geq \exp\{-c(d,\nu)\frac{t}{(\log t)^{2/d}}\}.$$

Probabilistically this situation corresponds to the following particle survival strategy. To survive among the traps until time t the particle starting from the origin travels to a pocket B_l with no obstacles and stays there for the reminder of time t. This scenario minimizes the cost in terms of probability $\mathbb{P}$ associated with the creation of a pocket free of obstacles. One has every right to think that there should be no reason for this behavior

of the Brownian particle to be typical. However, as it turns out, for large t the particle goes "extra mile" searching for some points x_l being "near minima" of a certain random functional and stays in their vicinity for the remainder of time t, cf. the so called *Pinning Effect* described in the following section.

By contrast the good survival strategy of a particle in the annealed case is to stay inside an obstacle free ball $B_t := B(0, \tilde{R}_t)$ of a certain radius $\tilde{R}_t = \tilde{R}t^{\frac{1}{d+2}}$ with $\tilde{R}$ to be determined later, where the chance of survival $\sim \exp\{-\lambda(B_t)t\} = \exp\{-\lambda(B(0,\tilde{R}))t^{\frac{d}{d+2}}\}$ is much larger than in the quenched case. The probability of an obstacle free hole of such a radius appearing in the Poisson cloud is $\sim e^{-\nu|B_t|}$. Hence asymptotically we obtain that

$$S_t \geq \exp\{-\left(\lambda_d/\tilde{R}^2 + \nu\omega_d\tilde{R}^d\right)t^{\frac{d}{d+2}}\}. \tag{12}$$

Optimizing over all $\tilde{R}$ we find that

$$S_t \geq \exp\{-\tilde{c}(d,\nu)t^{\frac{d}{d+2}}\}.$$

Here $\tilde{c}(d,\nu)$ is given by (8).

The optimal choice of $\tilde{R}$ in (12) is given by

$$\tilde{R}_0 = \left(\frac{2\lambda_d}{d\nu\omega_d}\right)^{1/(d+2)}. \tag{13}$$

Again as we shall see in the following section, cf. the so called *Confinement Property* this is not only the good strategy for survival but also the typical one in the annealed case.

3. PINNING EFFECT AND CONFINEMENT PROPERTY

3.1. Pinning Effect

We have described above a possible scenario under which the particle survives, in the quenched case, up to long time t. As we have already pointed out it is far from being clear that the strategy described there is typical. Furthermore it would be of some interest to find out which pockets of low potential are selected to stay in by the particle. To motivate the next result we present the following heuristic after Sznitman (1998).

We recall that according to what we have said in the previous section the presence of Poisson traps causes a localization effect for the particle expressed in the fact that most part of the support of any eigenfunction φ_i of H_ω is localized in a relatively small neighborhood of a certain point x_i and the function decays exponentially, say at the rate $\alpha|x - x_i|$ from its respective domain of localization. Thus the sequence $< \varphi_i, 1 >$, $i \geq 1$ appearing in (11) is almost constant of positive value. Using further (11) we can see that the asymptotic behavior of $S_{t,\omega}$ is governed by the term

$$e^{-\min_i(\alpha|x_i|+t\lambda_i)}.$$

Thus the particle should go by time t to some pocket of low potential located near a minimum of a certain random functional of the form

$$\alpha|x| + t\lambda_\omega(B(x, R_t)) \tag{14}$$

with R_t being some random localization scale. Here $\lambda_\omega(U)$ denotes the principal Dirichlet eigenvalue of H_ω in an open set U.

This heuristic picture has some truth in it as it is shown by Theorem 3 below. Before presenting the result let us first introduce some notation. Let us denote by

$$e_\lambda(x, y, \omega) = E_x \left[\exp\left\{ - \int_0^{H(y)} (\lambda + V(Z_s, \omega))ds \right\}, H(y) < +\infty \right]. \tag{15}$$

with $H(x)$ - the hitting time of the ball $\overline{B}(x) := \overline{B}(x, 1)$.

The strong Markov property of Brownian Motion implies that

$$e_\lambda(x, y, \omega) \geq e_\lambda(x, z, \omega) \inf_{\overline{B}(z)} e_\lambda(\cdot, y, \omega), \ \lambda \geq 0, \ x, y, z \in \mathbb{R}^d$$

which in turn guarantees the supermultiplicative property of $e_\lambda(x, y, \omega)$. By Liggett's ergodic theorem, cf. e.g. Liggett (1985), it is possible to obtain then the following, cf. Sznitman (1994),

THEOREM 2. — (The shape theorem.) *There exists a deterministic nontrivial norm $\alpha_\lambda(\cdot)$ on $\mathbb{R}^d$ satisfying for any $M > 0$*

$$\lim_{|x| \to +\infty} \sup_{0 \leq \lambda \leq M} \frac{1}{|x|} | - \log e_\lambda(0, x, \omega) - \alpha_\lambda(x)| = 0 \ \mathbb{P} \text{ a.s..} \tag{16}$$

Remarks.

1. The name appearing by the above theorem deserves a brief comment. The functions e_λ can be related to site "passage times" considered in the first time percolation theory cf. Kesten (1986), Hammersley-Welsh (1965). The theorem carries then some analogy to the shape theorem proven for the first passage percolation in Kesten (1986). The word "shape" refers here to the closed unit balls of the norm α_λ.
2. One can introduce also a random distance function given by

$$d_\lambda(x, y, \omega) := \max\{ - \inf_{\overline{B}(x)} \log e_\lambda(y, \cdot, \omega), - \inf_{\overline{B}(y)} \log e_\lambda(\cdot, x, \omega)\}. \tag{17}$$

It can be proven, cf. Proposition 5.2.2 of Sznitman (1998) that there exists a positive constant C depending only on the dimension d such that for all ω, $\lambda \geq 0$, $x, y \in \mathbb{R}^d$ with $|x - y| > 4$ one has

$$|d_\lambda(x, y, \omega) + \log e_\lambda(x, y, \omega)| \leq C(1 + F_\lambda(x) + F_\lambda(y)). \tag{18}$$

Here F_λ is a certain nonnegative function depending on ω, which is bounded if $d \geq 3$ or $\lambda > 0$ and otherwise of sublogarithmic growth $\mathbb{P}$ a.s., cf. Lemma 1.1 of Wüthrich

(1998). In consequence one can also claim the Shape Theorem with d_λ in place of e_λ. The theorem has actually even broader range of applicability and extends also to Green's function in place of e_λ, cf. Theorem 5.2.5 of Sznitman (1998) for more details.

One can introduce two random scales, $R_{t,\omega}$ satisfying

$$R_{t,\omega} = o(\exp\{(\log t)^{1-\chi}\}). \tag{19}$$

for some $\chi > 0$ and $S_\omega(t)$, which roughly speaking measures the distance of the locus of the farthest minimum of the random functional

$$F_t(x,\omega) := \alpha_0(x) + t\lambda_\omega(B(x, R_{t,\omega})) \tag{20}$$

from the origin. There exists such a $\chi > 0$ for which

$$\frac{t}{(\log t)^{\frac{2}{d}+\chi}} \geq S_\omega(t) \geq \exp\{(\log t)^{1-\chi}\}, \tag{21}$$

cf. Section 6.3 of Sznitman (1998).

The scale $R_{t,\omega}$ is larger than the macroscopic unit scale $(\log t)^{1/d}$ relevant to the quenched case but is much smaller than the scale $S_\omega(t)$ where the farthest near minima sites of $F_t(\cdot,\omega)$ occur. The importance of these scales is highlighted by Theorem 3 below. Before formulating the result let us introduce a bit more notation. For $\omega \in \Omega$ we define the event

$$C := [Z_\cdot \text{ comes before time } t \text{ within distance 1 of some } x \in \mathcal{L}_{c,t,\omega}$$

from which then it does not move any further than distance $R_{t,\omega}$ up to time t].

Here $\mathcal{L}_{c,t,\omega}$ stands for the skeleton of "near mininima" of $F_t(\cdot,\omega)$, i.e.

$$\mathcal{L}_{c,t,\omega} := [x \in \frac{1}{\sqrt{d}}\mathbb{Z}^d : F_t(\cdot,\omega) \leq \mu_{t,\omega} + cS_\omega(t)]$$

where $\mu_{t,\omega} := \min F_t(\cdot,\omega)$ and $c \in (0,1)$ is a constant. Since $\mu_{t,\omega} \sim c(d,\nu)\frac{t}{(\log t)^{2/d}}$ for $t \gg 1$, cf. (6.3.15) of Sznitman (1998) we can conclude thanks to (21) that $S_\omega(t) \ll \mu_{t,\omega}$ for large t. Hence we can see that the term "near minimum" is indeed well founded.

The statement of the pinning effect is then as follows.

THEOREM 3. — *For a sufficiently small $\chi > 0$ and $\mathbb{P}$ a.s. ω we have*

$$\lim_{t\uparrow+\infty} Q_{t,\omega}(C) = 1.$$

3.2. **Confinement Property**

We consider here only the hard obstacle case. The entire theory works also for soft obstacles (cf. Povel (1998a)). In this situation, as we recall, the particle gets instantaneously killed upon the contact with the obstacle field. It is convenient to rescale the entire problem so that the true obstacles have size $\varepsilon = t^{-\frac{1}{d+2}}$. The Brownian Motion, after rescaling, is given by $\varepsilon Z_{\frac{s}{\varepsilon^2}}$, $s \geq 0$. We can restrict our attention to those trajectories which do not leave, up to time $s := t^{\frac{d}{d+2}}$, the box $\mathcal{T} := [-t, t]^d$.

We have already mentioned that one possible survival strategy of the particle is to stay inside a ball B_0 centered at 0 of radius $\tilde{R}_0$. This ball minimizes the functional

$$\tilde{c}(d, \nu) := \inf_{U \text{ open}} (\nu|U| + \lambda(U)). \tag{22}$$

It can be proven, cf. Povel (1998) Proposition 1, that with large probability there exist random, open "clearing sets" $\mathcal{U}^{cl}_{\varepsilon,\omega}$ such that they are "almost" obstacle free and their complements are contained in the regions having high concentration of obstacles. In addition by virtue of Proposition 1 of Povel (1998) the clearing sets are "almost" optimal with respect to the variational problem (22) i.e. for some $\chi > 0$ we have

$$\mathbb{P}_\varepsilon \otimes P_0(\nu|\mathcal{U}^{cl}_{\varepsilon,\omega}| + \lambda(\mathcal{U}^{cl}_{\varepsilon,\omega}) \leq \tilde{c}(d, \nu) + \varepsilon^\chi \mid T_{\Theta_\varepsilon} > s) \to 1, \text{ when } t \uparrow +\infty$$

where T_{Θ_ε} is the exit time from the spatially rescaled obstacle free region Θ_ε (cf. (2)) and, as we recall, $s := t^{\frac{d}{d+2}}$. $\mathcal{U}^{cl}_{\varepsilon,\omega}$ can be defined as $\mathcal{T} \cap \mathcal{O}_\varepsilon(\omega) - D_\varepsilon(\omega)$, cf. Section 6 for the definition of $\mathcal{O}_\varepsilon(\omega)$ and $D_\varepsilon(\omega)$.

By Faber-Krahn inequality (cf. e.g. Chavel (1984)) we can conclude that

$$\nu|\mathcal{U}^{cl}_{\varepsilon,\omega}| + \lambda_d(\omega_d/|\mathcal{U}^{cl}_{\varepsilon,\omega}|)^{2/d} \leq \tilde{c}(d, \nu) + \varepsilon^\chi.$$

In consequence $\mathcal{U}^{cl}_{\varepsilon,\omega}$ must have both the volume and principal Dirichlet eigenvalue close to the corresponding quantities for the optimal ball. Using a strengthening of the Faber-Krahn result due to Hall (1992) one can choose a ball B_ω with radius $\approx \tilde{R}_0$ and such that $|\mathcal{U}^{cl}_{\varepsilon,\omega} - B_\omega|$ is small. This ball contains only a tiny portion of obstacles but it is surrounded by a dense forest of traps. Hence the rescaled particle must stay inside the ball if it is to survive up to time s. Along these lines, Sznitman (1991) for $d = 2$ and Povel (1998) have proven that:

THEOREM 4. — *For $d \geq 2$ there exists $\chi_1 > 0$, $1 > \chi_2 > 0$ such that for any ω there is a ball B^ω_0 with center in $B(0, (R_0 + \chi_1 t^{-\chi_2/(d+2)})t^{1/(d+2)})$ and a radius in $[R_0 t^{1/(d+2)}, (R_0 + \chi_1 t^{-\chi_2/(d+2)})t^{1/(d+2)}]$ for which*

$$\lim_{t\uparrow+\infty} Q_t[T_{B^\omega_0} > t] = 1.$$

Here $T_{B^\omega_0}$ denotes the exit time from the ball B^ω_0.

At this point let us also mention the following result due to Schmock (1990) in $d = 1$, Sznitman (1991) for $d = 2$ and Povel (1998) when $d \geq 3$.

THEOREM 5. — *As $t \uparrow +\infty$, the process $\left\{t^{-\frac{1}{d+2}} Z_{\tau t^{\frac{2}{d+2}}}\right\}_{\tau \geq 0}$ converges in law under Q_t to the mixture with weight $\varphi(x)/\int \varphi$ of the laws of Brownian motion starting from 0 and conditioned not to exit the ball $B(x, \tilde{R}_0)$. φ is the principal eigenfunction of $-\frac{1}{2}\Delta$ in $B(0, \tilde{R}_0)$.*

4. CROSSING BROWNIAN MOTION (CBM)

In this section we assume that the shape function W is rotationally invariant and the resulting random potential V is truncated on a certain level $M > 0$. We shall discuss the quenched case only. To our knowledge there are no corresponding results available for the annealed case.

The law of CBM on $\left(C(\mathbb{R}_+, \mathbb{R}^d), \mathcal{F}_{H(y)}\right)$ is defined by

$$d\hat{P}_x^y := \frac{1}{e_\lambda(x, y, \omega)} \exp\left\{-\int_0^{H(y)} (\lambda + V)(Z_s, \omega) ds\right\} 1_{[H(y)<+\infty]} dP_x,$$

where, as we recall $H(y)$ is the hitting time of the ball $\overline{B}(y)$, $\mathcal{F}_{H(y)}$ its σ-algebra and e_λ is given by (15).

The measure describes the behavior of the Brownian path in the potential $\lambda + V$ conditioned on the event that it performs a crossing from x to y. One of the questions of considerable physical interest related to CBM is an interplay between the strength of the transversal path fluctuation whose magnitude is supposed to be of a typical size $\sim |x-y|^\chi$ and the shape fluctuations of the random sphere given by the distance function defined in (17). The latter are assumed to be of order $|d_\lambda(0, y, \omega) - \alpha_\lambda(y)| \sim |y|^\xi$. ξ must belong to interval $[0, 1]$ by virtue of Theorem 2. It has been conjectured (cf. ibid) that the transverse fluctuations should satisfy the scaling identity

$$2\xi - 1 = \chi. \tag{23}$$

More specifically to measure the size of the aforementioned fluctuations let us define

$$\xi_1 := \inf\left\{\gamma \geq 0 : \lim_{|y|\uparrow+\infty} \mathbb{E}\left[\hat{P}_0^y[A(y, \gamma)]\right] = 1\right\},$$

where $A(y, \gamma)$ denotes the event that the trajectory starting at the origin and performing the crossing to $y \in \mathbb{R}^d$ stays inside a truncated cylinder of radius $|y|^\gamma$ centered on the axis l_y: αy, $\alpha \in \mathbb{R}$ and defined as

$$C(y, \gamma) := [z : \operatorname{dist}(z, l_y) \leq |y|^\gamma \text{ and } -|y|^\gamma \leq < z, y/|y| > \leq |y| + |y|^\gamma].$$

To measure the shape fluctuation of the distance function $d_\lambda(x, y, \omega)$ around its median $M_\lambda(x, y)$ we introduce

$$\chi^{(1)} := \inf\left\{\kappa \geq 0 : \lim_{|y|\uparrow+\infty} \mathbb{P}[|d_\lambda(0, y, \omega) - M_\lambda(0, y)| \leq |y|^\kappa] = 1\right\}$$

or alternatively

$$\chi^{(2)} := \inf\left\{\kappa \geq 0 : \lim_{r\uparrow+\infty} \mathbb{P}[\sup_{x\in\overline{B}(0,r)} |d_\lambda(0,x,\omega) - M_\lambda(0,x)| \leq r^\kappa] = 1\right\}. \tag{24}$$

The second coefficient measures the fluctuation size in a uniform manner thus obviously $\chi^{(1)} \leq \chi^{(2)}$. The following result is a corollary of Sznitman (1996):

THEOREM 6. — *For all $d \geq 2$ we have*

$$\chi^{(2)} \leq 1/2.$$

We can obtain also, cf. (23), that

$$2\xi_1 - 1 \leq \chi_2.$$

Indeed, in consequence of the rotational invariance and spatial homogeneity of the shape function W we get

$$M_\lambda(x,y) = M_\lambda(0, |y-x|e_1), \tag{25}$$

where $e_1 = (1, 0, \cdots, 0)$. From the Shape Theorem 2 we conclude easily that there exists a certain constant $C > 0$ for which

$$M_\lambda(0,x) \geq C|x| \qquad \text{for a sufficiently large } |x|. \tag{26}$$

Both here and in the sequel C shall stand for any generic constant independent of x, y.

We define C_y a finite set of points on $\partial C(y,\gamma)$ whose cardinality is less than or equal to $C|y|^d$ and such that $\partial C(y,\gamma) \subseteq \bigcup_{z\in C_y} \overline{B}(z)$.

Using the definition (24) of χ_2 we can write that

$$\mathbb{E}[\hat{P}_0^y[A(y,\gamma)^c]] \approx \tag{27}$$

$$\mathbb{E}\left\{A(\varepsilon), \hat{P}_0^y\left[\bigcup_{z\in C_y}[0 < H(z) < H(y) < +\infty]\right]\right\}$$

with

$$A(\varepsilon) := [\omega : \text{ the exponent } \kappa \text{ of the fluctuation of } d_\lambda \text{ does not exceed } \chi^{(2)} + \varepsilon].$$

Applying the strong Markov property of Wiener measure to stopping times $H(z) < H(y)$ we can estimate the probability $\hat{P}_0^y$ on the right hand side of the approximate equality (27) by

$$\frac{C}{e_\lambda(0,y,\omega)}|y|^d \sup_{z\in C_y} e_\lambda(0,z,\omega)e_\lambda(z,y,\omega). \tag{28}$$

Using a version of Harnack principle, cf. (5.2.22) of Sznitman (1998) and the definition (17) we can estimate (28) by

$$C|y|^d \sup_{z\in C_y} \exp\{d_\lambda(0,y,\omega) - d_\lambda(0,z,\omega) - d_\lambda(z,y,\omega)\} \leq \tag{29}$$

$$C|y|^d \sup_{z\in C_y} \exp\left\{M_\lambda(0,y) - M_\lambda(0,z) - M_\lambda(z,y) + 5|y|^{\chi^{(2)}+\varepsilon}\right\}.$$

The last line follows from the fact that $\omega \in A(\varepsilon)$.

Using relation (25) we can write that

(30)

$$M_\lambda(0,z) + M_\lambda(z,y) = M_\lambda(0,|z|e_1) + M_\lambda(|z|e_1,(|y-z|+|z|)e_1) \geq M_\lambda(0,(|y-z|+|z|)e_1).$$

The following lemma is a fairly straightforward consequence of the strong Markov property of Wiener measure and the aforementioned Harnack's principle, cf. Lemma 2.1 in Wüthrich (1998) for the proof.

LEMMA 1. — *There exists a constant $C > 0$ such that*

$$M_\lambda(0,(|y-z|+|z|)e_1) \geq M_\lambda(0,|y|e_1) + M_\lambda(|y|e_1,(|y-z|+|z|)e_1) - C|y|^{\chi^{(2)}+\varepsilon}, \tag{31}$$

for z, y, $y-z$ all sufficiently large in magnitude.

In consequence we can write that the right hand side of (29) can be estimated by

$$C|y|^d \sup_{z\in C_y} \exp\left\{M_\lambda(0,(|y-z|+|z|-|y|)e_1) + C|y|^{\chi^{(2)}+\varepsilon}\right\}. \tag{32}$$

Applying now an elementary inequality

$$|z| + |z-y| \geq |y| + C|y|^{2\gamma-1}$$

for all $z \in \partial C(y,\gamma)$ and $|y| > 1$ together with (26) we can see from (32) that

$$\mathbb{E}[\hat{P}_0^y[A(y,\gamma)^c]] \leq C\exp\{C_1|y|^{\chi_2+\varepsilon} - C_2|y|^{2\gamma-1}\},$$

for some positive constants C, C_1, C_2. The right hand side of the above inequality vanishes for large $|y|$ provided that $2\gamma - 1 > \chi_2 + \varepsilon$. Hence we have concluded the first part of the following result due to Wüthrich (1998).

THEOREM 7. — *For all $d \geq 2$ we have*

1) $\frac{\chi^{(2)}+1}{2} \geq \xi_1$
2) $\xi_1 \geq \frac{\chi_1+\chi_1/\chi_2}{2}$.

Remarks.

1. In order to obtain (23) we need the equality $\chi^{(1)} = \chi^{(2)}$, which is an open problem at this time.

2. It would be of considerable interest to prove a lower bound on $\chi^{(2)}$. It is believed that in two dimensions $\chi^{(2)} = \frac{1}{2}$.

5. INTEGRATED DENSITY OF STATES OF THE RANDOM SCHRÖDINGER OPERATOR

Let us consider a Random Schrödinger operator H_ω given by (10). By $N_{\Lambda,\omega}(\lambda)$ we denote the number of those Dirichlet eigenvalues of H_ω in Λ, counting multiplicities, which are less than or equal to λ. Here $\Lambda = (-l, l)^d$. We define then the thermodynamic limit

$$N(\lambda) := \lim_{l \uparrow +\infty} \frac{1}{|\Lambda|} N_{\Lambda,\omega}(\lambda)$$

whose existence $\mathbb{P}$ a.s. follows from the classical subadditive ergodic theorem, cf. Krengel (1985) or Carmona-LaCroix (1991). $N(\lambda)$ is a deterministic and increasing function called integrated density of states (IDS). In Quantum Physics $N(\lambda)$ describes a volume density of energy states lying below a fixed energy level λ for a Schrödinger operator with random Poissonian impurities. We are interested in determining the low energy level asymptotic of IDS, i.e. the behavior of $N(\lambda)$ when $\lambda \downarrow 0$. We define by $L(t)$ the Laplace transform of the measure $N(d\lambda)$ corresponding to IDS.

The classical Tauberian theorem (cf. e.g. Bingham et al. (1987), p. 254) tells us that the behavior of $N(\lambda)$ for small λ can be determined from the asymptotic behavior of $L(t)$ for large t. To characterize the latter we write

$$L(t) = \int_0^{+\infty} e^{-t\lambda} N(d\lambda) = \lim_{l \uparrow +\infty} \frac{1}{|\Lambda|} \int_0^{+\infty} e^{-t\lambda} N_{\Lambda,\omega}(d\lambda), \tag{33}$$

where $N_{\Lambda,\omega}(d\lambda)$ is the appropriate random measure associated with $N_{\Lambda,\omega}(\lambda)$. Using the trace formula and ergodic theorem we can represent the left hand side of (33) as

$$\mathbb{E} r_\omega(t, 0, 0) \tag{34}$$

with $r_\omega(t, x, y)$ the random kernel associated with an $L^2(R^d)$ strongly continuous semigroup

$$\left(e^{tH_\omega} f\right)(x) := E_x \left[f(Z_t) \exp\left\{ -\int_0^t V(Z_s, \omega) ds \right\} \right] = \int_{R^d} f(y) r_\omega(t, x, y) dy.$$

Using the expectation with respect to the Brownian Bridge measure $P_t^{x,y}$ we obtain that

$$r_\omega(t, x, y) = (2\pi t)^{-d/2} p(t, x, y) E_t^{x,y} \exp\left\{ -\int_0^t V(Z_s, \omega) ds \right\}.$$

The asymptotic of $L(t)$ and S_t for large t can be shown to be identical.

Indeed, thanks to (34) we can write that

$$L(t) = \mathbb{E} \int_{R^d} r_\omega(t-1, 0, y) r_\omega(1, y, 0) dy \leq c S_{t-1}.$$

On the other hand using again (34) and denoting by U^a the a-neighborhood of U we obtain

$$L(t) = \frac{1}{|U|}\mathbb{E}\int_U r(t,x,x)dx \geq \frac{1}{|U|}\mathbb{E}[e^{-\lambda_\omega(R^d)t},\ \text{no obstacle falls into } U^a] = \frac{1}{|U|}e^{-\lambda(U)t-\nu|U^a|}$$

and the last quantity is bounded below by

$$e^{-\tilde{c}(d,\nu)t^{\frac{d}{d+2}}(1+O(1))}.$$

for a suitable choice of U. Here $\tilde{c}(d,\nu)$ is given by (8). We recall here that $B(0,a)$ contains the support of W and $\lambda_\omega(\cdot)$, $\lambda(\cdot)$ stand for the local Dirichlet eigenvalues of H_ω, $-\frac{1}{2}\Delta$ respectively. Applying now Theorem 1 we obtain the following.

THEOREM 8. — *The integrated density of states satisfies*

$$N(\lambda) = e^{-\nu|B(0,\sqrt{\frac{\lambda_d}{\lambda}})|(1+o(1))} \text{ for } \lambda \downarrow 0. \tag{35}$$

Remarks.

1. It can be proven, cf. Pastur et al. (1992), that in the case when the shape function W decays at infinity at the rate $\sim \frac{C}{|x|^\alpha}$, $|x| \gg 1$ (the so called classical regime), with $\alpha \in (d, d+2)$, the tails of IDS behave like $N(\lambda) \sim \exp\{-C(d,\nu,W)\lambda^{-\frac{d}{\alpha-d}}\}$, $\lambda \ll 1$. The constant $C(d,\nu,W)$ depends then on the shape function which is in sharp contrast with the situation discussed in Theorem 8 where the corresponding constant depended only on the field intensity ν and the dimension d, cf. (35). The result of the theorem can be extended to cover also the case of the shape function having $\alpha > d+2$ (nonclassical regime).
2. Let us also mention here that an analogous result to Theorem 8 has been obtained in the dimension two for the magnetic Schrödinger operator

$$H_\omega := (-i\nabla + A(\cdot))^2 + V(\cdot,\omega)$$

in Erdösz (1998). Here $A(x_1,x_2) = (-\frac{Bx_2}{2}, \frac{Bx_1}{2})$. Erdösz has proved that the integrated density of states defined for the bottom of the L^2-spectrum of H_ω, i.e. $\frac{B}{2}$, is given by

$$N(B/2+\lambda) \sim \lambda^{-\frac{2\pi\nu}{B}}$$

for small $\lambda > 0$.

6. THE METHOD OF ENLARGEMENT OF OBSTACLES (MEO)

The problems described in Sections 2, 3.1 and 5 highlight the importance of controls over the local Dirichlet eigenvalues $\lambda_\omega(U)$ of the open set U. Actually the upper bounds of the eigenvalues, which correspond to the lower bounds on the quenched or annealed

asymptotics, are quite easy to derive thanks to the monotonicity property of local eigenvalues with respect to their respective domains. This explains relative simplicity of our "derivation" of the lower bounds on both quenched and annealed asymptotics given in Section 2. It is considerably more difficult to obtain lower bounds on $\lambda_\omega(U)$. They allow to control S_t, $S_{t,\omega}$ of Section 2 from above, thanks to the estimates of the type

$$0 \sup_x E_x \left[\exp\left\{ -\int_0^t V(Z_s,\omega)ds \right\}, T_U > t \right] \leq c[1+(\lambda_\omega(U)t)^{d/2}]e^{-\lambda_\omega(U)t}, \tag{36}$$

cf. (3.1.9) of Sznitman (1998).

In the first step we rescale for convenience space and time by $x' = \varepsilon x$, $t' = \varepsilon^2 t$ with ε corresponding to the inverse of the typical spatial scale of the problem on hand, e.g. the heuristics of Section 2 suggest that ε equals $(\log t)^{-1/d}$ in the quenched and $t^{-1/(d+2)}$ in annealed case. The true obstacles in the rescaled situation are of size $\sim \varepsilon$. We wish to obtain lower bounds of $\lambda_\omega^\varepsilon(U)$ -the Dirichlet principal eigenvalue of the scaled Schrödinger operator $-\frac{1}{2}\Delta + \frac{1}{\varepsilon}V(\frac{\cdot}{\varepsilon})$ consider in a region U.

In U we can distinguish three disjoint subregions corresponding to various degrees of the concentration of obstacles. In the close neighborhood of the first region there is a high density of obstacles so that the particle entering it "feels" them very quickly and thus gets trapped in short time. This fact allows to "solidify" those obstacles by imposing there Dirichlet zero boundary condition, or equivalently to remove this subregion of U from the consideration without significantly influencing the local eigenvalue.

In the second region the concentration of obstacles is too small to perform the kind of "surgery" just described without significant distortion of the magnitude of the local eigenvalue. However it can be proven, cf. the "Volume Control" Theorem 11 below, that this part of U occupies asymptotically vanishing fraction of volume, as $\varepsilon \downarrow 0$, so we can discard it entirely without significantly changing the upper bound on the eigenvalue coming from the Faber-Krahn isoperimetric theorem, cf. Chavel (1984).

The remaining third part of region U receives no obstacles.

Let us explain now the method by getting an upper estimate of the annealed asymptotic of S_t. For simplicity we consider here only the case when $d \geq 3$. First we wish to construct a coarse grained picture of the region where $V > 0$. Unfortunately the use of a fine scale, say of order of magnitude of obstacles, will result in too high a number of possible region shapes. In effect we will not be able to obtain any meaningful control of the survival probability. This fact provides motivation to enlarge the true obstacles, which are of size ε, say to the size ε^γ, $0 < \gamma < 1$ in order to lower the combinatorial complexity of the problem. The *density set* $D_\varepsilon(\omega)$ corresponding to the first region is defined then as the set having a lot of obstacles in its immediate neighborhood.

More specifically for any fixed integer $L > 1$ we consider the lattice $\mathbb{L}_{n_\gamma(\varepsilon)} := \frac{1}{L^{n_\gamma(\varepsilon)}}\mathbb{Z}^d$, with $n_\gamma(\varepsilon) := [\gamma \frac{\log \frac{1}{\varepsilon}}{\log L}]$. We define $D_\varepsilon(\omega)$ as the union of those boxes C from the lattice for

which

$$\sum_{n_\alpha(\varepsilon)<k<n_\gamma(\varepsilon)} L^{k(d-2)}\mathrm{cap}_k \geq \delta(n_\gamma(\varepsilon) - n_\alpha(\varepsilon)). \tag{37}$$

Here $\delta > 0$, $0 < \alpha < \gamma$ are some parameters and cap_k denotes the capacity of the obstacles lying in the box from the lattice $\mathbb{L}_k$ containing C. The parameter $n_\alpha(\varepsilon)$ determines the "testing spatial scale" for condition (37) and equals $n_\alpha(\varepsilon) := [\alpha \frac{\log \frac{1}{\varepsilon}}{\log L}]$. Notice that $\varepsilon^\alpha \gg \varepsilon^\gamma \gg \varepsilon$ when $\varepsilon \downarrow 0$.

We have then two crucial spectral controls, cf. Theorems 4.2.3 and 4.2.6 of Sznitman (1998).

THEOREM 9. — *There exists a positive constant $C(d, W)$ depending only on the dimension d and the shape function W such that for any $M > 0$ and $\varrho \in (0, \delta C \frac{\gamma-\alpha}{(d+2)\log L})$ we have*

$$\limsup_{\varepsilon\downarrow 0} \sup_{U,\omega} \varepsilon^{-\varrho} \left(\lambda^\varepsilon_\omega(U - \overline{D}_\varepsilon(\omega)) \wedge M - \lambda^\varepsilon_\omega(U) \wedge M\right) = 0$$

where the sup *is taken over any $\omega \in \Omega$ and open $U \subseteq \mathbb{R}^d$.*

To formulate the second result we need a notion of a clearing set. For a given parameter $r > 0$ we declare a box C of the unit lattice $\mathbb{Z}^d$ to be a *clearing box* if

$$|C - D_\varepsilon(\omega)| \geq r^d.$$

If otherwise holds we call C a *forest box.*

Let $\mathcal{A}_\varepsilon(\omega)$ be the union of all clearing boxes. Suppose that $\mathcal{O}_\varepsilon(\omega)$ is the R-neighborhood of $\mathcal{A}_\varepsilon(\omega)$. We have then the following.

THEOREM 10. — *Let $M > 0$. There exist $C(d) > 0$, $C'(d, M) > 1$, $r_0(d, M) \in (0, \frac{1}{4})$ such that*

$$\limsup_{\varepsilon\downarrow 0} \widetilde{\sup} \exp\left\{C\left[\frac{R}{4r}\right]\right\} (\lambda^\varepsilon_\omega(U \cap \mathcal{O}_\varepsilon(\omega)) \wedge M - \lambda^\varepsilon_\omega(U) \wedge M) = 0$$

provided that $\varepsilon^\alpha < r < r_0$, $R/(4r) > C'$. $\widetilde{\sup}$ is taken over all configurations ω and open sets U.

For the second region $B_\varepsilon(\omega)$ we take boxes of size $\sim \varepsilon^\beta$ for some $\gamma < \beta < 1$ from the appropriate lattice $\mathbb{L}_{n_\beta(\varepsilon)}$ which are contained in the complement of density boxes and receive some points of the obstacle cloud. We have then

THEOREM 11. — *Let $d \geq 3$. There exists $L_0(d)$ such that for $L \geq L_0$ we have $\delta_0(d, L) > 0$ for which for any $\delta < \delta_0$*

$$\limsup_{\varepsilon\downarrow 0} \sup_C \varepsilon^{-\kappa_0} |B_\varepsilon(\omega) \cap C| < +\infty$$

where sup *is taken over all unit size boxes C, $\kappa_0 > 0$ depends on all the parameters involved except for ε and ω.*

The third and last region is $R^d - (D_\varepsilon(\omega) \cup B_\varepsilon(\omega))$. Of course, as we have already mentioned before, it contains no points of the cloud. Observe that for any unit box C the sets $D_\varepsilon(\omega) \cap C$ and $B_\varepsilon(\omega) \cap C$ can take no more than $2^{\varepsilon^{-d\gamma}}$ and $2^{\varepsilon^{-d\beta}}$ different possible shapes respectively. This should be contrasted with the number of possible shapes of the true obstacle set which is of order of complexity $2^{\varepsilon^{-d}} \gg 2^{\varepsilon^{-d\beta}} \gg 2^{\varepsilon^{-d\gamma}}$, when $\varepsilon \downarrow 0$.

We apply the above framework to obtain an upper bound for the annealed norming constant S_t. We start by choosing a large box $\mathcal{T} := (-t, t)^d$.

Let us define now the "essential" event E consisting of those clouds for which $\lambda_\omega^\varepsilon(\mathcal{T}) \leq 2\tilde{c}(d, \nu)$ and $\mathcal{A}_\varepsilon(\omega) \cap \mathcal{T}$ consists of at most n_0 clearing boxes where

$$\frac{\nu}{2} n_0 r^d \geq 2\tilde{c}(d, \nu).$$

We can easily observe that

$$\mathbb{P}_\varepsilon[E^c] \leq (4t)^{dn_0} \mathbb{P}_\varepsilon[[[0,1)^d \subseteq \mathcal{A}_\varepsilon(\omega)]^{n_0}. \tag{38}$$

From the definition of the clearing set and Theorem 11 we can conclude that

$$\mathbb{P}_\varepsilon[[0,1)^d \subseteq \mathcal{A}_\varepsilon(\omega)] = \mathbb{P}_\varepsilon[|[0,1)^d - D_\varepsilon(\omega)| \geq r^d] \leq$$

$$\mathbb{P}_\varepsilon[|[0,1)^d - (D_\varepsilon(\omega) \cup B_\varepsilon(\omega))| \geq r^d - \varepsilon^{\kappa_0}].$$

Since the set $[0,1)^d - (D_\varepsilon(\omega) \cup B_\varepsilon(\omega))$ contains no point of the cloud we can estimate the left hand side of (38) by

$$2^{2\varepsilon^{-d\beta}} \exp\left\{-\nu \varepsilon^{-d} r^d\right\} \leq \exp\left\{-\frac{\nu}{2} \varepsilon^{-d} r^d\right\}$$

for sufficiently small ε (or equivalently for sufficiently large t).

From (38) and (6) we obtain that S_t is, up to term of order $e^{-2\tilde{c}(d,\nu)s}$ $(s = t^{\frac{d}{d+2}})$, less than or equal to

$$\mathbb{E}_\mathfrak{e} \otimes E_0 \left[\exp\left\{ -\int_0^s V_\varepsilon(Z_u) du \right\}, E \right].$$

We are ready now to apply Theorems 9, 10 and 11. Set $M := 2\tilde{c}(d, \nu)$. Define

$$U(\omega) := \mathcal{T} \cap \mathcal{O}_\varepsilon(\omega) - D_\varepsilon(\omega)$$

$$V(\omega) := \mathcal{T} \cap \mathcal{O}_\varepsilon(\omega) - (D_\varepsilon(\omega) \cup B_\varepsilon(\omega)).$$

From Theorem 11 we obtain that

$$|U(\omega)| \leq |V(\omega)| + |\mathcal{T} \cap \mathcal{O}_\varepsilon(\omega)| \varepsilon^{\kappa_0} \leq |V(\omega)| + (2R+1)^d n_0 \varepsilon^{\kappa_0}. \tag{39}$$

Observe that the number of possible shapes of $U(\omega)$, $V(\omega)$ cannot exceed $2^{C\varepsilon^{-d\beta}}$, for some $C > 0$. For a given deterministic U_0 and V_0 we consider the event $\mathcal{G}_{U_0,V_0}$ consisting of those ω for which $U(\omega) = U_0$, $V(\omega) = V_0$. Thanks to (36) we can estimate as follows

$$\mathbb{E}_\epsilon \otimes E_0 \left[\mathcal{G}_{U_0,V_0} \cap E, \exp\left\{ -\int_0^s V_\varepsilon(Z_u)du \right\} \right] \leq$$

$$c(d)(1+(2\tilde{c}s)^{d/2})\mathbb{E}_\epsilon \otimes E_0 \left[\mathcal{G}_{U_0,V_0} \cap E, \exp\left\{-(\lambda^\varepsilon_\omega(\mathcal{T}) \wedge M)s\right\}\right] \leq$$

$$c(d)(1+(2\tilde{c}s)^{d/2})\mathbb{P}_\varepsilon[\mathcal{G}_{U_0,V_0}] \exp\left\{-\lambda^\varepsilon_\omega(\mathcal{T} \cap \mathcal{O}_\varepsilon(\omega)) \wedge Ms\right\}.$$

In the last line we used Theorem 10. Applying Theorem 9 we obtain then from (39) that

$$S_t \leq 2^{C\varepsilon^{-d\beta}} \exp\{-\inf_{\text{all admissible } U_0} (\lambda(U_0) + \nu|U_0|) \wedge Mt^{\frac{d}{d+2}} + \text{ smaller order terms}\} \leq$$

$$\exp\{-\tilde{c}(d,\nu)t^{\frac{d}{d+2}} + \text{ smaller order terms}\}.$$

This proves the upper bound in part 2) of Theorem 1.

7. OPEN PROBLEMS

Let us start our review of some of the open questions with the problem of defining a "right" notion of random geodesics for Brownian motion moving among random traps. We have already mentioned in Section 4 that $d_\lambda(\cdot,\cdot,\omega)$ defined by (17) is a metric. However it is not a geodesic type of a distance. One can verify that

$$d_\lambda(x,y,\omega) + d_\lambda(y,z,\omega) = d_\lambda(x,z,\omega)$$

implies that at least two of x, y, z must be identical. In this context it is worthwhile to mention the existence of a continuous analogue of the first passage time geodesic given by

$$\varrho_\lambda(x,y,\omega) := \inf_{\gamma \in \mathcal{P}(x,y,1)} \left\{ \int_0^1 \sqrt{2(\lambda+V)(\gamma_s,\omega)}|\gamma'(s)|ds \right\}.$$

Here $\mathcal{P}(x,y,1)$ is the set of Lipschitz paths γ leading from x to y in time 1. It can be verified that the analogue of the Shape Theorem 2 holds for ϱ_λ with the limiting deterministic norm denoted by $\mu_\lambda(\cdot)$. On the other hand one can obtain, cf. Wüthrich (1998a) that for any Euclidean norm unit vector e

$$\mu_\lambda(e) = \lim_{\beta\uparrow+\infty} \frac{\alpha_{\beta,\lambda}(e)}{\sqrt{\beta}}$$

where $\alpha_{\beta,\lambda}$ is the norm resulting from a shape theorem for the quantities

$$e_{\beta,\lambda}(x,y,\omega) := E_x\left[\exp\left\{-\beta\int_0^{H(y)}(\lambda+V(Z_s,\omega))ds\right\}, H(y)<+\infty\right], \tag{40}$$

cf. Theorem 2. One can view then $-\log e_{\beta,\lambda}(x,y,\omega)/\sqrt{\beta}$ for large β as some sort of distance function. In determining the path integral one uses, in light of the fact that the potential becomes infinite for $\beta \uparrow +\infty$, only the most "optimal" paths which so to say perform the fastest possible crossings. The proof that the appropriate limit exists is not available yet. It leads also to the question of determining random "pseudo" geodesics, i.e. tubes of width small in comparison with the crossing distance which contains with large probability the paths of the CBM. This problem has also connections with better formulation of the scaling identity (23) in terms of the intrinsic geodesic instead of Euclidean cylinders.

In the context of the results presented in Section 5, it is worthwhile to mention that the approach to the asymptotic of IDS via Tauberian theory is not entirely satisfactory. A direct proof of the Lifshitz tail could be constructed via counting the number of energy states lying below certain λ for the Dirichlet problem in a sufficiently large "a priori" box. It is clear that the typical (unit) size associated with the problem is of order of magnitude $1/\sqrt{\lambda}$. Using MEO it should be possible to show that one can in fact put a Dirichlet boundary on the forest part of the box without distorting significantly the behavior of IDS for small λ. The expected result would probably yield asymptotic of IDS up to higher order terms, cf. Theorem 8.

Finally let us also mention that the random scales $S_\omega(t)$ and $R_{t,\omega}$ introduced in the subsection discussing the pinning effect are rather poorly understood at this time. They characterize the location of the farthest near minimum of the random functional (20) and the size of the pocket in which a pinned Brownian Motion is located after arriving in the vicinity of a relevant near minimum. It would be desirable to get better asymptotic of those scales than those provided by (19) and (21). In addition it would be of considerable interest to obtain some better understanding of the magnitude of times at which the particle arrives at the pinning location. The results in that direction concern mostly the one dimensional case. A reader is encouraged to consult Section 6.5 of Sznitman (1998) for details.

REFERENCES

[1] Bingham, N. H., Goldie, C. M., Teugels, J. L. - *Regular variations*, vol. 27, Encyclopedia of Mathemathics and its applications, Cambridge Univ. Press (1987).

[2] Carmona, R., LaCroix, J., I. - *Spectral theory of random Schrödinger operator*, Probability and Application, Birkhäuser, Basel (1991).

[3] Chavel, I. - *Eigenvalues in Riemannian Geometry*, Academic Press, New York (1984).

[4] Donsker, M., Varadhan, S. R. S. - *Asymptotic for the Wiener sausage*, Comm. Pure Appl. Math. **28** (1975), 525-565.

[5] Erdösz, L. - *Lifshitz tail in a magnetic field: the non-classical regime.* Preprint (1998).

[6] Grassberger, P., Procaccia, I. - *The long time properties of diffusion in a medium with static traps*, J. Chem. Phys. **77** (1982), 6281-6285.

[7] Hammersley, J. M., Welsh, D. J. A. - *First passage percolation, subadditive processes, stochastic networks and generalised renewal theory.* In *Bernoulli, Bayes, Laplace Anniversary volume*, L.M. Le Cam and J. Neyman editors, Springer, New York (1965).

[8] Kac, M., Luttinger, J. M. - *Bose Einstein condensation in the presence of impurities* I, II, J. Math. Phys. **14** (1973), 1626; **15** (1974), 183-186.

[9] Kesten, H. - *Aspects of first passage percolation*, vol 1180, Springer Lecture Notes in Math., New York (1986).

[10] Krengel, U. - *Ergodic Theorems*, de Gruyter, Berlin-New York (1985).

[11] Krug, J., Spohn, H. - *Kinetic roughening of growing surfaces.* In *Solids far from equilibrium*, C. Godréche, Ed. Cambridge Univ. Press (1991), 479-582.

[12] Lifshitz, I. M. - *Energy spectrum structure and quantum states of disordered condensed systems*, Sov. Phys. Uspekhi **7** (1965), 549-573.

[13] Liggett, T. - *Interacting particle systems*, Springer, New York (1985).

[14] Pastur, L., Figotin, A. - *Spectra of random and almost periodic operators*, New York, Springer (1992).

[15] Povel, T. - *Confinement of Brownian Motion among poissonian obstacles in* $\mathbb{R}^d$, $d \geq 3$ (1998), to appear in Prob. Theory and Rel. Fields.

[16] Povel, T. - Private communication (1998a).

[17] Schmock, U. - *Convergence of the normalized one dimensional Wiener sausage path measure to a mixture of Brownian taboo processes*, Stochastics **29** (1990), 171-183.

[18] Smoluchowski, M. V. - *Versuch einer mathemathischen Theorie der koagulationskinetischen Lösungen*, Z. Phys. Chemie **92** (1918), 129-168.

[19] Spitzer, F. - *Principles of Random Walk*, van Nostrand, Princeton (1964).

[20] Sznitman, A. S. - *Lifshitz tail and Wiener sausage*, I, II, J. Funct. Anal. **94** (1990), 223-246, 247-272.

[21] Sznitman, A. S. - *On the confinement property of Brownian Motion among Poisson obstacles*, Comm. Pure Appl. Math. **44** (1991), 1137-1170.

[22] Sznitman, A. S. - *Shape theorem, Lyapunov exponents and large deviation for Brownian motion in a poissonian potential*, Comm. Pure Appl. Math. **47** (1994), 1655-1688.

[23] Sznitman, A. S. - *Quenched critical large deviations for Brownian motion in a poissonian potential*, Journ. Funct. Anal. **131** (1995), 54-77

[24] Sznitman, A. S. - *Annealed Lyapunov exponents and large deviations in a poissonian potential I,II*, Ann. Scient. Éc. Norm. Sup. **28** (1995a,b), 345-370, 371-390.

[25] Sznitman, A. S. - *Distance fluctuations and Lyapunov exponents*, Ann. of Prob. **24** (1996), 1507-1530

[26] Sznitman, A. S. - *Brownian Motion, Obstacles and Random Media*, Springer Monographs in Mathemathics, Springer-Verlag Berlin, Heidelberg, New York (1998).

[27] Wüthrich, M. - *Scaling identity for crossing Brownian Motion in a poissonian potential* (1998), to appear in Prob. Theory and Rel. Fields.

[28] Wüthrich, M. - *Geodesics and crossing Brownian motion in a soft poissonian potential* (1998a), to appear in Ann. Inst. H. Poincaré.

[29] Zerner, M. -*Lyapunov exponents and quenched large deviation for multidimensional random walk in a random environment* (1997), to appear in Ann. of Prob.

Tomasz KOMOROWSKI

Institute of Mathemathics
Maria Curie - Skłodowska University
Lublin - Poland
E-mail : komorow@golem.umcs.lublin.pl

Astérisque

STEFANO MARMI

Chaotic behaviour in the solar system

Astérisque, tome 266 (2000), Séminaire Bourbaki,
exp. n° 854, p. 113-136

<http://www.numdam.org/item?id=SB_1998-1999__41__113_0>

© Société mathématique de France, 2000, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (http://smf4.emath.fr/Publications/Asterisque/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Séminaire BOURBAKI Novembre 1998
51ème année, 1998-99, n° 854

CHAOTIC BEHAVIOUR IN THE SOLAR SYSTEM
[following J. LASKAR]

by **Stefano MARMI**

0. INTRODUCTION

I. Newton certainly believed that the Solar System is topologically unstable. In his view the perturbations among the planets were strong enough to destroy the stability of the Solar System. He even made the hypothesis that God controls the instabilities so as to insure the existence of the Solar System: "but it is not to be conceived that mere mechanical causes could give birth to so many regular motions .... This most beautiful system of the sun, planets, and comets, could only proceed from the counsel and dominion of an intelligent powerful Being" [N, p. 544]. The problem of Solar System stability was (and for many aspects still is) a real one: Halley was able to show, by analyzing the Chaldean observations transmitted by Ptolemy, that Saturn was moving away from the Sun while Jupiter was getting closer. A crude extrapolation leads to a possible collision in 6 million years (Myrs) in the past.

From a mathematical point of view arguments supporting the long–time stability of the orbits of the planets were given by Lagrange, Laplace and Poisson who proved the absence of secular evolution (polynomial increase in time) of the semi–major axis of the planets up to third order in the planetary masses.

On the contrary the researches of Poincaré [P] and Birkhoff [B] showed that instabilities might occur in the dynamics of the planets and that the phase space must have a quite complicated structure.

In the course of the year 1954, Kolmogorov [K] stated his famous theorem of persistence of quasiperiodic motions in near to integrable Hamiltonian systems, and first suggested that the picture may be twofold: stability in the sense of measure theory conjugated with topological instability. Arnold moreover proved [Ar1] that bounded orbits have positive measure in the planar three–body problem and claimed

that the same result must be true for the n–body problem (provided that the masses of the planets are sufficiently – unrealistically – small). He also first proved in the general context of Hamiltonian systems with many degrees of freedom the existence of orbits which drift (or diffuse) along resonances so as to change by a finite amount their action [Ar2][1], even if this process is very slow [Ne]. This reinforced the belief that "the time after which chaos manifests itself under a sufficiently small perturbation of the initial state is large in comparison with the time of existence of the Solar System" [Ar3, p.82].

Following Herman [He2] one can legitimately ask the following question: *If one of the masses* $m_0 = 1$ *and all the other masses* $m_j \ll 1$ *are sufficiently small, are there wandering domains*[2] *in any neighborhood of fixed distinct circular orbits around the mass* m_0 *and moving in the same direction in a plane?*

Quite recently some progress has been made in the heuristic understanding of the dynamics of the planets of the Solar System, due largely to the help provided by computers but also due to a better understanding of the underlying dynamics, resulting from the great progress in the overall field of Dynamical Systems. Modern computers allow extensive analytic calculations and numerical integrations of realistic models over very long times, even if the shortness of the step–size needed for the computation has for many years limited the investigation to the outer planets of the Solar System (Jupiter, Saturn, Uranus, Neptune and Pluto) [CMN], [SW]. Indeed, the faster the orbital movement of the planet is, the shorter is the step–size required (from approximately 40 days for Jupiter to 12 hours for Mercury). As a result, until 1991 the only available numerical integration of a realistic model of the full Solar System was spanning only 44 centuries. For this reason the analytical approach, which makes use of perturbation theory, is needed.

J. Laskar replaced the full Newtonian equations of the motion by the so–called *secular system* introduced by Lagrange where the fast angular variables are eliminated. This system, instead of giving the fast motion of the planets in space, describes the slow deformation of the planets' orbits. In this way Laskar reduced the number of degrees of freedom of the system and achieved an impressing reduction of the step–size required (most of the computational time in traditional numerical integrations is actually spent in the numerical solution of Kepler's problem). In fact he was able to

[1] "Thus, even if the motion of a planet or an asteroid is regular, an arbitrarily small perturbation of the initial state is sufficient to make it chaotic" [Ar3, p.82].

[2] i.e. an open set V and $t_0 > 0$ such that $f^t(V) \cap V = \emptyset$ for all $t > t_0$, where f^t denotes the Hamiltonian flow.

use a step–size of 500 years. The numerical integration of this system shows that *the inner Solar System* (Mercury, Venus, Earth and Mars) *is chaotic* with a Lyapounov time of 5 Myrs. This measures the rate of the exponential growth of the distance in phase space between the orbits of two points initially close [Yo]. As a consequence it is not possible to compute ephemeris for the position of the Earth over 100 Myrs: an error of 15 meters in the initial position of the Earth may grow to an error of 150 million kilometers (i.e. its present distance to the Sun) after 100 Myrs. This kind of strong instability could even result in the escape of Mercury in 3.5 billion years (Gyrs). The deformations of the planets' orbits is responsible for an external forcing on the Hamiltonian describing the evolution of the obliquity of each planet. The obliquity is the angle between the equator and the orbital plane. Laskar shows that it can undergo dramatic variations on a time scale very short in geological terms.

In what follows we will describe these results and the ideas underlying Laskar's approach, mainly coming from the theory of Hamiltonian systems and classical perturbation theory. We will also briefly discuss the technique of the numerical analysis of the fundamental frequency developed by Laskar to study the mixed phase space structure of quasi–integrable Hamiltonian systems.

The style of this exposition will be quite informal, partially because most of the results reported here lack a rigorous justification (and sometimes even a good mathematical formulation). In the last section we will try to formulate some open problems inspired by Laskar's work.

Acknowledgments. In the preparation of this review I have extensively used references [Ar1], [L1995a], [L1996] and some unpublished seminar notes of M. Herman [He3]. I have also benefited a lot from many discussions with A. Chenciner, J. Laskar, D. Sauzin and J.–C. Yoccoz.

0.1. Hamiltonian systems, integrable systems, quasiperiodic motions [Ar4]

Usually in mechanics the equations of the motion of a conservative system with phase space $M = T^*N$ (here the configuration space N is an f–dimensional riemannian manifold) are given in Hamiltonian form: $\dot{p}_i = -\frac{\partial H}{\partial q_i}$, $\dot{q}_i = \frac{\partial H}{\partial p_i}$, $1 \leq i \leq f$. Here the "generalized coordinates" q_i and their "conjugate momenta" p_i are a system of local canonical coordinates in M and $H : M \to \mathbb{R}$ is smooth (the Hamiltonian of the system). Note that in many problems arising from celestial mechanics the flow is not complete due to the unavoidable occurence of collisions (see [Ch2] for a recent review). The symplectic form on M is $\omega = \sum_{i=1}^{f} dp_i \wedge dq_i$ and maintains this

expression in all canonical systems of coordinates (they form an atlas by Darboux' theorem). Two functions $F, G \in \mathcal{C}^\infty(M, \mathbb{R})$ are *in involution* if their Poisson bracket $\{F, G\} = 0$, i. e. when their Hamiltonian flows commute.

An important extension of the Hamiltonian formalism is obtained considering time–dependent Hamiltonian functions $H : M \times \mathbb{R} \to \mathbb{R}$. These are especially useful, as we will see, for modeling non–isolated systems, i. e. mechanical systems under the action of some external forcing.

An especially interesting case is provided by the manifold $\mathbb{R}^f \times \mathbb{T}^f$ which can be identified with the cotangent bundle of the f–dimensional torus $\mathbb{T}^f = \mathbb{R}^f/(2\pi\mathbb{Z})^f$. This manifold has a natural symplectic structure defined by the closed 2–form $\omega = \sum_{i=1}^f dJ_i \wedge d\vartheta_i$ where $(J_1, \dots J_f, \vartheta_1, \dots \vartheta_f)$ is a point on $\mathbb{R}^f \times \mathbb{T}^f$. Let U denote an open connected subset of $\mathbb{R}^f$. Whenever an Hamiltonian system can be reduced by a symplectic change of coordinates to a function $H : U \times \mathbb{T}^f \to \mathbb{R}$ which *does not depend on the angular variables* ϑ one says that the system is *completely canonically integrable* and the variables J are called *action variables*. Note that in this case Hamilton's equations take the particularly simple form $\dot{J}_i = -\frac{\partial H}{\partial \vartheta_i} = 0$, $\dot{\vartheta}_i = \frac{\partial H}{\partial J_i}$, $i = 1, \dots, f$. Let $\nu_i(J) = \frac{\partial H}{\partial J_i}$, $i = 1, \dots, f$. The associated flow $t \mapsto (J(t) \equiv J(0), \vartheta(t) = \vartheta(0) + t\nu(J(0)))$, $t \in \mathbb{R}$, is *linear* and leaves invariant the f–dimensional torus $J = J(0)$. The motion is therefore bounded and *quasiperiodic* if the $\mathbb{Z}$–module $\{k \in \mathbb{Z}^f, k \cdot \nu(J(0)) = 0\}$ has dimension at most $f - 2$. Otherwise the motion is periodic. The Hamiltonian is *non degenerate* (i.e. satisfies the "twist condition") if $\det\left(\frac{\partial^2 H}{\partial J_i \partial J_k}(J)\right) \neq 0$ for all $J \in U$, thus the "frequency map"

$$J \mapsto \nu(J) = \frac{\partial H}{\partial J}(J) \in \mathbb{R}^f \tag{0.1}$$

is a local diffeomorphism. This condition is generic, but in many applications and especially in the two–body problem (see below) the Hamiltonian H is *properly degenerate*: $\det\left(\frac{\partial^2 H}{\partial J_i \partial J_k}(J)\right) = 0$ for all $J \in U$. In this case the linear flow on the invariant tori can be described by only $f_0 < f$ frequencies, where $f_0 = \operatorname{rank}\left(\frac{\partial^2 H}{\partial J_i \partial J_k}\right)$, by suitably choosing new action–angle coordinates $J' = (A^T)^{-1} J$, $\vartheta' = A\vartheta$, $A \in \mathrm{GL}(f, \mathbb{Z})$.

0.2. The two–body problem and action–angle variables. If the mutual attraction of the planets is neglected, each planet is attracted only by the Sun. This leads to the two–body problem whose Hamiltonian $\mathcal{H}_0 : T^*(\mathbb{R}^3 \setminus \{\mathbf{0}\}) \mapsto \mathbb{R}$ in the center of

mass frame is[3]

$$\mathcal{H}_0(\mathbf{p},\mathbf{q}) = \frac{1}{2\mu}\|\mathbf{p}\|^2 - \frac{m_0 m}{\|\mathbf{q}\|} \tag{0.2}$$

where m_0 is the mass of the Sun, m is the mass of the planet and $\mu = m_0 m/(m_0+m)$ is the reduced mass of the system. $\mathcal{H}_0$ is called the *Kepler Hamiltonian.* For negative energy, the solutions are ellipses with one focus at the origin (i.e. the center of mass). These are called *Keplerian orbits.* The shape and the position of the ellipse in space are determined from the knowledge of the semi-major axis a, the eccentricity e, the angle of inclination i of its plane w.r.t. the horizontal plane $q_3 = 0$, the argument of the perihelion ω and the longitude of the ascending node Ω (fig. 1). The position of the planet along the ellipse is determined by the mean anomaly l which is proportional to the area swept by the position vector $\mathbf{q}$ of the planet starting from the perihelion.

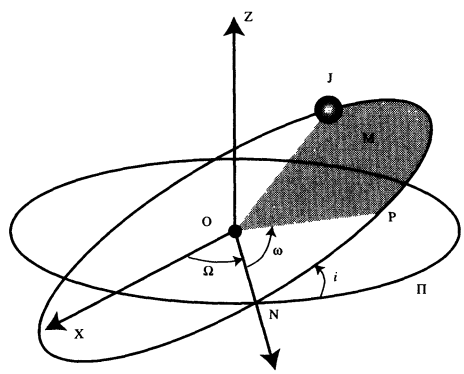


Fig. 1. Elliptical elements of a Keplerian orbit (from [L1995a])

The system admits 5 *independent* first integrals: the total energy H, the three components of the angular momentum $\mathbf{q}\wedge\mathbf{p}$ and one of the components of the Laplace vector $\mathbf{A} = \frac{1}{m}\mathbf{p}\wedge(\mathbf{q}\wedge\mathbf{p}) - \frac{m_0 m}{\|\mathbf{q}\|}\mathbf{q}$. Among these integrals one can choose three integrals in involution and construct the canonical transformation to action–angle variables (see [P2], Ch. III and [Ch1]). The other two integrals are responsible for the proper complete degeneration of the Kepler problem: for all initial conditions the orbit is periodic and the period depends only on the energy. A set of action–angle

[3] We have set the universal gravitational constant $= 1$.

coordinates $(L, G, \Theta, l, g, \theta) \in \mathbb{R}^2_+ \times \mathbb{R} \times \mathbb{T}^3$ such that two of the three frequencies of the linear flow on the invariant tori vanish is given by *Delaunay variables.* They are related to the orbital elements as follows:

$$L = \mu\sqrt{(m_0+m)a}\ ,\quad G = L\sqrt{1-e^2}\ ,\quad \Theta = G\cos i\ ,\quad l\ ,\quad g = \omega\ ,\quad \theta = \Omega\ .$$

Note that G is the modulus of angular momentum $\mathbf{q} \wedge \mathbf{p}$, thus Θ is its projection along the q_3–axis. One has the obvious limitation $|\Theta| \leq G$. The new Hamiltonian reads $\mathcal{H}_0 = -\frac{\mu^3(m_0+m)^2}{2L^2}$ thus $\nu_g = \frac{\partial \mathcal{H}_0}{\partial G} = \nu_\theta = \frac{\partial \mathcal{H}_0}{\partial \Theta} = 0$. The Delaunay variables are not suitable for the description of the orbits of the planets of the Solar System since they are singular at circular orbits ($e = 0$, thus $L = G$ and the argument of the perihelion g is not defined) and at horizontal orbits ($i = 0$ or $i = \pi$, thus $G = \Theta$ and the longitude of the ascending node θ is not defined). But all the planets of the Solar System have almost circular orbits (with the exception of Mercury and Mars) and small inclinations (see Table 1).

Planet	ϖ	Ω	i	e	λ	a
Mercury	77.4561	48.3309	7.0050	0.205632	252.2509	0.387104
Venus	131.5637	76.6799	3.3947	0.006772	181.9798	0.723307
Earth	102.9373			0.016709	1000.4664	1.000012
Mars	336.0602	49.5581	1.8497	0.093401	355.4333	1.523711
Jupiter	374.3313	100.4644	1.3033	0.048495	34.3515	5.210216
Saturn	93.0568	113.6655	2.4889	0.055509	50.0775	9.538070
Uranus	173.0052	74.0060	0.7732	0.046296	314.0550	19.183302
Neptune	48.1237	131.7841	1.7700	0.008989	304.3487	30.055144

TABLE 1: Orbital elements of the planets of the Solar System (without Pluto): ϖ, Ω, i and λ are given in degrees, a is given in Astronomical Units (1 A.U. $= 1.5 \cdot 10^8$ km).

This problem is solved by introducing a new set of action–angle variables $(\Lambda, H, Z, \lambda, h, \zeta) \in \mathbb{R}^3 \times \mathbb{T}^3$: $\Lambda = L$, $H = L - G$, $Z = G - \Theta$, $\lambda = l + g + \theta$, $h = -g - \theta$, $\zeta = -\theta$ (the variable λ is called the mean longitude, $-h$ is the longitude of the perihelion) then considering the couples (H, h) and (Z, ζ) as polar symplectic coordinates:

$$(0.3)\quad \xi_1 = \sqrt{2H}\cos h\ ,\quad \eta_1 = \sqrt{2H}\sin h\ ,\qquad \xi_2 = \sqrt{2Z}\cos\zeta\ ,\quad \eta_2 = \sqrt{2Z}\sin\zeta\ .$$

The variables $(\Lambda, \xi, \lambda, \eta) \in \mathbb{R}_+ \times \mathbb{R}^2 \times \mathbb{T}^1 \times \mathbb{R}^2$ are called *Poincaré variables.* They are well defined also in the case of circular ($H = 0$) or horizontal ($Z = 0$) orbits. The

new Hamiltonian reads $\mathcal{H}_0 = -\frac{\mu^3(m_0+m)^2}{2\Lambda^2}$ thus $\dot{\lambda} = \frac{\mu^3(m_0+m)^2}{\Lambda^3}$ whereas Λ, ξ and η are constant. The relation between Poincaré variables and the original momentum–position $(\mathbf{p}, \mathbf{q})$ variables is much more subtle and will not be discussed here (see [P2], Ch. III, [L1989a]). Note however that Λ is proportional to $\sqrt{a}$,

$$\sqrt{\xi_1^2 + \eta_1^2} \simeq \sqrt{\Lambda} e(1 + \mathcal{O}(e^2)) \,, \quad \sqrt{\xi_2^2 + \eta_2^2} \simeq \sqrt{\Lambda} i(1 + \mathcal{O}(i^2) + \mathcal{O}(e^2)) \,.$$

Astronomers actually use the non–canonical coordinates given by the orbital elements themselves.

0.3. KAM theory, Nekhoroshev theorem, Arnol'd diffusion. Following Poincaré, the *fundamental problem of dynamics* is the study of quasi–integrable Hamiltonian systems, i.e. Hamiltonians $\mathcal{H} : U \times \mathbb{T}^f \mapsto \mathbb{R}$ (smooth or analytic) of the form

$$\mathcal{H}(J, \chi) = h_0(J) + \varepsilon h_1(J, \chi) \,, \tag{0.4}$$

where ε is a small real parameter. As we will see in Sect. 1.1, the Solar System Hamiltonian has this form with $f = 3n$, where n is the number of the planets. h_0 is equal to the sum of n independent Kepler Hamiltonians and ε is of the order of the planetary masses. The Keplerian orbits give an n–dimensional torus invariant for the flow associated to h_0.

Most results have been obtained under the assumption that the unperturbed Hamiltonian h_0 is non degenerate. The general picture is provided by KAM [Bo, Y1] and Nekhoroshev [Ne, Lo] theorems: if ε is sufficiently small, most initial conditions (w.r.t. Lebesgue measure) lie on invariant f–dimensional Lagrangian tori carrying quasiperiodic motions with diophantine frequencies. The action variables corresponding to these KAM orbits will remain ε–close to their initial values for all times. The complement of this set is open and dense and it is connected if $f \geq 3$. It contains a connected ($f \geq 3$) web $\mathcal{R}$ of resonant zones corresponding to $\mathbb{Z}^f$–linearly dependent frequencies: $\cup_{k \in \mathbb{Z}^f} \{J \in U \,, \nu_0(J) \cdot k = 0\} \times \mathbb{T}^f$. Motion along these resonances cannot be excluded, resulting in a variation of $\mathcal{O}(1)$ of the actions in a finite time[4]. But if the Hamiltonian is *analytic* and h_0 is *steep* in the sense of [Ne] (for example convex) then this variation is very slow: it takes a time at least $\mathcal{O}\left(\exp\left(\frac{1}{\varepsilon^a}\right)\right)$ to change the

[4] It is conjectured [Ar4, p. 189] that generically quasi–integrable Hamiltonians with more than two degrees of freedom are topologically unstable.

actions by $\mathcal{O}(\varepsilon^b)$, where a and b are two positive constants. Moreover each invariant torus $\mathcal{T}$ has a neighborhood filled in with trajectories which remain close to it for an even longer time [MG]. Indeed assume that h_0 is convex and the frequencies ν of the linear flow on $\mathcal{T}$ satisfy a diophantine condition (see [Y1]) of exponent $\tau \geq n-1$, i.e. there exists $\gamma > 0$ such that

$$|k \cdot \nu| \geq \gamma |k|^{-\tau} , \quad \forall k \in \mathbb{Z}^f \setminus \{0\} , \tag{0.5}$$

where $|k| = |k_1| + \ldots + |k_f|$. Then all trajectories starting at a distance of order $\rho < \rho^*$ from $\mathcal{T}$ will remain close to it for a time $\mathcal{O}\left(\exp\left(\exp\left(\frac{\rho^*}{\rho}\right)^{1/(\tau+1)}\right)\right)$.

Unfortunately the proper degeneration of the Kepler Hamiltonian makes the application of these results to the Solar System problematic. One can introduce the secular system of Lagrange and use the perturbation to remove the degeneracy. However, this leads to a small twist which is of the order of ε and the application of KAM theory to the Solar System Hamiltonian is a very delicate task [He3].

A much simpler but still important case is given by the problem of stability of Lagrange's equilateral equilibrium solutions in the restricted three–body problem [Li1,Li2,Pol,SM,Wi]. In the Sun–Jupiter situation computer–assisted proofs can be made so accurate to prove the practical stability of some Trojan asteroids for 10 Gyrs [GDFGS,GS].

1. FREQUENCY MAP ANALYSIS

1.1. The frequency map analysis. Let $N_0 \subset \mathbb{R}^f$ be the image of U under the frequency map $\nu = \frac{\partial h_0}{\partial J} : U \to N_0$ and assume that this map is a diffeomorphism. One of the consequences of KAM theorem is the existence, for sufficiently small values of ε, of a Cantor set $N_\varepsilon \subset N_0$ of values of the frequencies ν for which the Hamiltonian system (0.4) has smooth invariant tori with linear flow with frequencies ν. This Cantor set corresponds to diophantine frequencies (0.5) with $\tau > n-1$ and $\gamma = \gamma_\varepsilon$. Moreover there exists a diffeomorphism

$$F_\varepsilon : N_0 \times \mathbb{T}^f \to U \times \mathbb{T}^f , \quad (\nu, \varphi) \mapsto (J, \vartheta) \tag{1.1}$$

which on $N_\varepsilon \times \mathbb{T}^f$ "straightens out all these invariant tori at the same time" ([Pö], p. 655), i.e. transforms Hamilton's equations of motion of (0.4) into $\dot{\varphi} = \nu$, $\dot{\nu} = 0$. This diffeomorphism is ε–close to $\left(\frac{\partial h_0}{\partial J}, \mathrm{id}_{\mathbb{T}^n}\right)$ and restricted to $N_\varepsilon \times \mathbb{T}^f$ is smooth in the sense of Whitney w.r.t. the first factor and analytic w.r.t. the second (if the

Hamiltonian (0.4) is analytic). Let us fix $\vartheta = 0$ and consider the diffeomorphism $\mathcal{F}_\varepsilon : N_0 \to U$ induced by (1.1). If the initial conditions $(J(0), 0)$ lie on a KAM torus with frequency $\nu \in N_\varepsilon$ each component z_r of $(J_1 e^{i\vartheta_1}, \dots, J_f e^{i\vartheta_f})$ will be an analytic quasiperiodic function of time ε–close to $J_r(0)e^{i\nu_r t}$. Given $t_0, T \in \mathbb{R}$, $T > 0$, the *frequency map analysis* [L1993, L1995b] is the numerical construction of an approximate inverse $\mathcal{F}^{-1}_{\varepsilon,T,t_0}$ of $\mathcal{F}_\varepsilon$ *on the whole actions space* U from the datum of $\{z_r(t), r = 1, \dots, f, \ t \in [t_0, t_0 + 2T]\}$. As $T \to +\infty$, this approximate inverse converges to $\mathcal{F}^{-1}_\varepsilon$ on the set of KAM tori.

1.2. Construction of the approximate frequency map on the set of KAM tori. Assume for simplicity that $t_0 = -T$ and that z_r has a Fourier series of the form $z_r(t) = e^{i\nu_r t} + \sum_{k \in \mathbb{Z}^f, k \neq e_r} \hat{z}_k e^{ik\cdot\nu t}$ with $|\hat{z}_k| < 1$ (here e_r is the r–th canonical basis vector of $\mathbb{Z}^f$, thus $e_r \cdot \nu = \nu_r$). Laskar looks for a quasiperiodic approximation to z_r of the form of a finite sum $\zeta_r^{(N)}(t) = \sum_{s=1}^N \hat{\zeta}_r^{(s)} e^{i\mu_r^{(s)} t}$, where the coefficients $\hat{\zeta}_r^{(s)}$ have decreasing amplitude with s, $\mu_r^{(1)} \simeq \nu_r$ and $\mu_r^{(s)} \simeq k^{(r,s)} \cdot \nu$ for some suitably chosen $k^{(r,s)} \in \mathbb{Z}^f$, $s = 2, \dots N$.

This quasiperiodic approximation is chosen in the following way.

1. Consider the weighted L^2 scalar product

$$\langle z, w \rangle_{T,p} = \frac{1}{2T} \int_{-T}^{T} z(t)\overline{w(t)} \chi_p\left(\frac{t}{T}\right) dt \,,$$

where z and w are smooth quasiperiodic functions. Here χ_p is the so–called Hanning window filter: $\chi_p(t) = \frac{2^p (p!)^2}{(2p)!}(1 + \cos \pi t)^p$ but one may consider more general situations. Let $\hat{\chi}_p(x) = \langle e^{ixt}, 1 \rangle_{1,p} = \frac{(-1)^p \pi^{2p} (p!)^2 \sin x}{x(x^2-\pi^2)\dots(x^2-p^2\pi^2)}$. Note that $\hat{\chi}_p(x) = 1 - c_p x^2 + \mathcal{O}(x^3)$ as $x \to 0$, for some $c_p > 0$, $\hat{\chi}_p(x) = \mathcal{O}(x^{-2p-1})$ as $x \to \infty$ as well as all its derivatives.

2. Let $y = (\nu_r - x)T$. Then

$$f(y, T) = \langle z_r, e^{ixt} \rangle_{T,p} = \hat{\chi}_p(y) + \sum_{k \in \mathbb{Z}^f, k \neq e_r} \hat{z}_k \hat{\chi}_p(y + (k \cdot \nu - \nu_r)T)$$

verifies $\lim_{T \to +\infty} [f(y, T) - \hat{\chi}_p(y)] = 0$ uniformly w.r.t. $y \in K$, where K is any compact subset of $\mathbb{R}$. Applying the implicit function theorem to the function $(y, T) \mapsto \frac{\partial}{\partial y} |f(y, T)|^2$ in a neighborhood of the point $(0, +\infty)$ one obtains the existence of a unique value x_r of x which maximizes $x \mapsto |\langle z_r, e^{ixt} \rangle_{T,p}|$ for sufficiently large T.

3. Let $\mu_r^{(1)} = x_r$, $r = 1, \ldots, f$, and define $\mathcal{F}^{-1}_{\varepsilon,T,t_0}(J(0)) = \mu = (\mu_1^{(1)}, \ldots, \mu_r^{(1)})$.

The pointwise convergence of $\mathcal{F}^{-1}_{\varepsilon,T,t_0}$ restricted to the set of KAM tori to $\mathcal{F}_\varepsilon^{-1}$ follows from the following asymptotic estimate proved in [L1995b]: there exists $c_{p,\varepsilon} > 0$ such that $|\nu_r - \mu_r^{(1)}| \leq c_{p,\varepsilon}(p!)^2 T^{-2p-2}$. This error term actually is the first term of an asymptotic series in powers of $\frac{1}{T}$ which is Gevrey. Applying the classical method of summation of the series to the smallest term one may choose $p = p(T)$ so as to obtain an exponentially small error term but for the time being this seems to be more mathematically entertaining than numerically useful.

To obtain $\zeta_r^{(N)}$ one iterates the above procedure: once the first periodic approximation $e^{i\mu_r^{(1)}t}$ is obtained, $\hat{\zeta}_r^{(1)}$ is computed by orthogonal projection and the process is started again on the remaining part $z_{r,1} = z_r - \hat{\zeta}_r^{(1)} e^{i\mu_r^{(1)}t}$ of the function.[5]

1.3. Applications of the frequency map analysis. Outside the set of KAM tori one can still apply the method described above but the result lacks a rigorous justification. This is mainly due to the fact that the dynamics "between" KAM tori is far from being understood. Keeping T and the initial condition $(J(0), \vartheta(0))$ fixed one computes $\mu(t) = \mathcal{F}^{-1}_{\varepsilon,T,t}(J(0))$ for different values of t. The time–dependence of μ is then used by Laskar to measure the "diffusion" of the orbit since, as we have seen, for KAM orbits μ is constant.

This method was first introduced by Laskar [L1990] in his study of the Solar System in order to estimate the size of the chaotic zones. Outside the domain of celestial mechanics it has been applied to many different dynamical situations. For exact area preserving twist maps, applying the classical theory of Birkhoff for KAM curves [He1] (these must be graphs of Lipschitz functions) one can derive a very practical criterion for the non existence of KAM curves [LFC]. The frequency map analysis has also been succesfully applied to the study of realistic models of particle accelerators [LRo].

[5] Note the analogy between this procedure and the epicycloid theory of Ptolemy where one adds up uniform circular motions to represent the motion of the planets: see [St] for a very beautiful exposition.

2. CHAOTIC BEHAVIOUR IN THE SOLAR SYSTEM

2.1. Planetary theory in Poincaré's canonical heliocentric coordinates.[6] Let us consider $n+1$ bodies with masses $m_0, \dots m_n$. Let O denote the center of mass of the system and $u_i \in \mathbb{R}^3$ the coordinate of the i–th body in a fixed barycentric reference frame. The conjugate momenta are $v_i = m_i \dot{u}_i$ and the Hamiltonian of the system is simply given by $\mathcal{H}(v,u) = \sum_{i=0}^{n} \frac{v_i^2}{2m_i} - \sum_{0 \le i < j \le n} \frac{m_i m_j}{\|u_i - u_j\|}$. Clearly $\mathcal{H}$ is defined on $T^*(\mathbb{R}^{3(n+1)} \setminus \Delta)$ where $\Delta = \{(u_0, \dots, u_n) \in \mathbb{R}^{3(n+1)}\ ,\ \exists i \neq j\ ,\ u_i = u_j\}$. Heliocentric canonical coordinates are induced by the linear transformation in configuration space $q_0 = u_0\ ,\quad q_i = u_i - u_0\ ,\quad 1 \le i \le n$ thus the conjugate momenta become $p_0 = \sum_{i=0}^{n} v_i\ ,\quad p_i = v_i\ ,\quad 1 \le i \le n$. By Noether's Theorem the classical first integrals of the system can be deduced from the symmetries of the Hamiltonian: the invariance with respect to uniform translations in configuration space implies the conservation of the total momentum p_0, thus it is not restrictive to fix $p_0 = 0$ (reduction of the center of mass). Conservation of energy is a consequence of the independence of the Hamiltonian from time and its invariance with respect to rotations of $\mathbb{R}^3$ implies the conservation of total angular momentum $C = \sum_{i=1}^{n} q_i \wedge p_i$. After reduction of the center of mass the number of degrees of freedom becomes $3n$ and in canonical heliocentric coordinates the Hamiltonian is given by $\mathcal{H}(p,q) = \mathcal{H}_0(p,q) + \mathcal{H}_1(p,q)$ with

$$
\begin{aligned}
\mathcal{H}_0(p,q) &= T_0(p) + V_0(q) = \sum_{i=1}^{n} \frac{\|p_i\|^2}{2\mu_i} - \sum_{i=1}^{n} \frac{m_0 m_i}{\|q_i\|} \\
\mathcal{H}_1(p,q) &= T_1(p) + V_1(q) = \sum_{0<i<j\le n} \frac{p_i \cdot p_j}{m_0} - \sum_{0<i<j\le n} \frac{m_i m_j}{\|q_i - q_j\|}
\end{aligned}
\tag{2.1}
$$

where $\frac{1}{\mu_i} = \frac{1}{m_i} + \frac{1}{m_0}$. $\mathcal{H}_0$ is the Hamiltonian of a collection of n independent Kepler problems each relative to a planet of mass μ_i attracted by a fixed point at the origin of mass $m_0 + m_i$. In the case considered by Laskar $n = 8$ (he did not consider Pluto whose mass is negligible), the phase space has 48 dimensions and is foliated with 8–dimensional tori invariant for the flow associated to the Kepler Hamiltonian $\mathcal{H}_0$.

[6] Actually Laskar in his numerical computations did use the same set of non canonical heliocentric coordinates traditionally used by astronomers since Laplace which allowed him comparison with the semianalytical theory of the Solar System of Bretagnon [Br]. This is irrelevant for our purposes. In the words of Poincaré [P, tome II, p. 37] "Les équations où s'introduisent les crochets de Lagrange prennent ainsi une forme en apparence plus compliquée. Mais cette différence n'a rien d'essentiel".

Choosing units $m_0 = 1$, $m_i = \varepsilon M_i$, $i = 1, \dots n$, where $\varepsilon \leq 10^{-3}$ for the Solar System, the conformally symplectic coordinate change $p = \varepsilon P$, $q = Q$, $\mathcal{H} = \varepsilon \hat{\mathcal{H}}$ shows that $\hat{\mathcal{H}}$ has the form (0.4) of a quasi integrable properly degenerate Hamiltonian system and that $\mathcal{H}_1$ is formally of the first order (w.r.t. $\mathcal{H}_0$) in the planetary masses.

In order to study the long term behaviour of the planetary orbits it is convenient to introduce for each planet Poincaré's variables $(\Lambda_i, \xi_i, \lambda_i, \eta_i)$, $i = 1, \dots n$, $\xi_i = (\xi_{1i}, \xi_{2i}) \in \mathbb{R}^2$, $\eta_i = (\eta_{1i}, \eta_{2i}) \in \mathbb{R}^2$. The Hamiltonian becomes

$$\mathcal{H}(\Lambda, \lambda, \xi, \eta) = \mathcal{H}_0(\Lambda) + \mathcal{H}_1(\Lambda, \lambda, \xi, \eta) \tag{2.2}$$

where $(\Lambda, \xi, \lambda, \eta) \in \mathbb{R}^n_+ \times \mathbb{R}^{2n} \times \mathbb{T}^n \times \mathbb{R}^{2n}$ and $\mathcal{H}_0(\Lambda) = -\sum_{i=1}^n \frac{\mu_i^3 (m_0+m_i)^2}{2\Lambda_i^2}$. The three components of the total angular momentum are

$$C_1 - iC_2 = \sum_{k=1}^n \sqrt{G_k^2 - \Theta_k^2}\, e^{i\zeta_k}\ , \quad C_3 = \sum_{k=1}^n \Lambda_k - \frac{1}{2}(\xi_{1k}^2 + \xi_{2k}^2 + \eta_{1k}^2 + \eta_{2k}^2) \tag{2.3}$$

where $G_k^2 - \Theta_k^2 = \frac{1}{2}(\xi_{2k}^2 + \eta_{2k}^2)\left(\Lambda_k - \xi_{1k}^2 - \eta_{1k}^2 - \frac{1}{2}(\xi_{2k}^2 + \eta_{2k}^2)\right)$.

2.2. Averaging and the secular system. In order to study the behaviour of the orbits in a neighborhood of a fixed Keplerian orbit we translate $\Lambda = \Lambda_0 + \tilde{\Lambda}$. Neglecting constant terms and the terms of order at least three in $\tilde{\Lambda}$ we obtain a new Hamiltonian $\tilde{\mathcal{H}} = \tilde{\mathcal{H}}_0 + \tilde{\mathcal{H}}_1$, where $\tilde{\mathcal{H}}_0 = N \cdot \tilde{\Lambda}$ and

$$\tilde{\mathcal{H}}_1(\tilde{\Lambda}, \lambda, \xi, \eta) = \mathcal{H}_1(\Lambda_0, \lambda, \xi, \eta) + f_1(\Lambda_0, \lambda, \xi, \eta) \cdot \tilde{\Lambda} + f_2(\Lambda_0, \lambda, \xi, \eta)\tilde{\Lambda} \cdot \tilde{\Lambda}\ .$$

In these formulas $N_j = \frac{\partial \mathcal{H}_0}{\partial \Lambda_j}(\Lambda_0) = \frac{\mu^3 (m_0+m_j)^2}{\Lambda_{0j}^3}$ is the mean motion of the j–th reference Keplerian orbit, $(f_1)_i = \frac{\partial \mathcal{H}_1}{\partial \Lambda_i}$, $(f_2)_{ij} = \frac{1}{2}\left(\frac{\partial^2 \mathcal{H}_0}{\partial \Lambda_i \Lambda_j} + \frac{\partial^2 \mathcal{H}_1}{\partial \Lambda_i \Lambda_j}\right)$, $i, j = 1, \dots, n$. Note that $\tilde{\mathcal{H}}$ is already an approximation of (2.2) since we have neglected all the terms $\mathcal{O}(\tilde{\Lambda}^3)$. Following Poincaré [P] we now look for a canonical transformation to new variables $(\Lambda', \lambda', \xi', \eta')$ *formally* close to the identity such that the new Hamiltonian $\mathcal{H}'$ does not depend on λ' when terms of the third order in the planetary masses are neglected

$$\mathcal{H}'(\Lambda', \xi, \eta) = \mathcal{H}'_0(\Lambda') + \mathcal{H}'_1(\Lambda', \xi', \eta') + \mathcal{H}'_2(\Lambda', \xi', \eta') \tag{2.4}$$

where $\mathcal{H}'_1$ and $\mathcal{H}'_2$ are of the first and of the second order respectively in the planetary masses. This "partial averaging procedure" (or resonant normal form) is a standard

technique in the study of quasi integrable Hamiltonian systems (see, for example, [Ar4], Ch. 5) and leads to the *secular system* studied by Laskar [L1992]. A first order calculation easily shows that $\mathcal{H}'_0(\Lambda') = N \cdot \Lambda'$ and $\mathcal{H}'_1(\Lambda', \xi', \eta') = \langle \tilde{\mathcal{H}}_1 \rangle = \int_{\mathbb{T}^n} \tilde{\mathcal{H}}_1(\Lambda', \lambda, \xi', \eta') d\lambda$. The frequencies vector $N \in \mathbb{R}^n$ must be non resonant (i.e. linearly independent on $\mathbb{Z}^n$) and this imposes a condition on Λ_0. For the same reason the coordinate change, as well as the Hamiltonian $\mathcal{H}'$, cannot be globally defined in the phase space.

Clearly the new variables Λ' are first integrals of the Hamiltonian flow of $\mathcal{H}'$. They are thus omitted from the numerical integration. Of course, this does not necessarily prove that the original variables Λ (i.e. the actual major semiaxes) are constant in time. Indeed this is not true. Since $\dot{\Lambda} = \{\Lambda, \mathcal{H}_1\}$, at the first order in the planetary masses one has $\{\Lambda, \mathcal{H}_1\} = \{\Lambda', \mathcal{H}_1\}$, whose average w.r.t. λ is zero. This is the classical theorem of Laplace [Lp1, Lp2] and Lagrange [Lg] asserting the absence of secular terms in the time evolution of the major semiaxes at first order in the masses. Poisson [Po] generalized this result to the second order whereas it is false at order three [Ha].

frequency ("/year)	**period (years)**
$g_1 = 5.596$	231000
$g_2 = 7.456$	174000
$g_3 = 17.365$	74600
$g_4 = 17.916$	72300
$g_5 = 4.249$	305000
$g_6 = 28.221$	45900
$g_7 = 3.089$	419000
$g_8 = 0.667$	1940000
$s_1 = -5.618$	230000
$s_2 = -7.080$	183000
$s_3 = -18.851$	68700
$s_4 = -17.748$	73000
$s_5 = 0.000$	
$s_6 = -26.330$	49200
$s_7 = -3.005$	431000
$s_8 = -0.692$	1870000

TABLE 2: Secular frequencies of the solar system (J. Laskar [L1996, p. 167])

$\mathcal{H}'_1$ is an even function of (ξ', η') and has an equilibrium position at $\xi' = \eta' = 0$. By a linear symplectic coordinate change $(\xi'', \eta'') = S(\xi', \eta')$ its quadratic part $\mathcal{H}'_{12}(\xi', \eta')$ can be written

$$\mathcal{H}'_{12}(\xi'', \eta'') = \frac{1}{2}\sum_{k=1}^{n}[g_k((\xi''_{1k})^2 + (\eta''_{1k})^2) + s_k((\xi''_{2k})^2 + (\eta''_{2k})^2)] \tag{2.5}$$

where the *secular frequencies* $g_k > 0$, $k = 1, \dots n$, $s_5 = 0$, $s_k < 0$, $k \neq 5$ (see Table 2).

The vanishing of s_5 is due to the existence of the total angular momentum integral (2.3). Note that the labeling of the secular frequencies does not correspond to the labeling of the planets since the matrix S is not diagonal. Thus the choice of $s_5 = 0$ is conventional. The solutions of Hamilton's equations associated to (2.5) are

$$\xi''_{1k}(t) + i\eta''_{1k}(t) = \sum_{j=1}^{n} \alpha_{kj} e^{ig_j t}\ ,\ \ \xi''_{2k}(t) + i\eta''_{2k}(t) = \sum_{j=1}^{n} \beta_{kj} e^{is_j t}\ , \tag{2.6}$$

thus the secular frequencies describe the (small amplitude) quasi periodic time–dependence (in this integrable approximation) of the inclinations and the eccentricity of the planets (fig. 2).

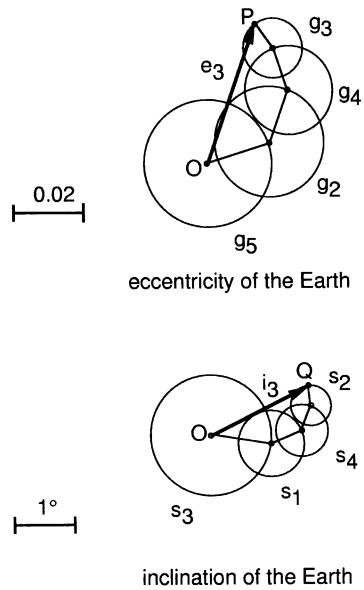


Fig. 2. Quasiperiodic variations of the Earth eccentricity and inclination according to Laplace and Lagrange (from [L1995a])

The Keplerian ellipses are no longer fixed: they are subject to a double precessionary motion (precession of the perihelion and of the line of nodes) with periods ranging from 45000 to 1940000 years. The quite long time scale of these phenomena has allowed Laskar to use a time–step of 500 years in the numerical integrations of Hamilton's equations associated to $\mathcal{H}'$.

2.3. Planetary evolution over millions and billions of years. The actual secular systems numerically integrated by Laskar [L1989, L1990] is accurate up to second order w.r.t. the masses and up to degree 5 in the eccentricities and inclinations (i.e. is obtained retaining polynomial terms of degree up to 6 w.r.t. the variables ξ, η in the Hamiltonian (2.4)). It contains about 150000 polynomial terms and the secular effects of general relativity and the Moon represent a few terms which have been added to the secular system. Its accuracy has also been tested comparing it with a direct numerical integration of the full Solar System which spans 6 Mys [LQT].

The first main result obtained by Laskar is that the whole solar system, and more specifically the inner Solar System (Mercury, Venus, Earth and Mars) is chaotic with a very short Lyapounov time of 5 Myrs. Since the typical time scale of the Laplace–Lagrange integrable secular system (2.5) is about 50000 years this Lyapounov time corresponds just to about 100 "periods". This has the conceptually striking consequence that it is practically impossible to predict the motion of the planets beyond 100 Myrs. This chaotic behaviour is essentially due to the presence of two secular resonances among the planets: $\theta = 2(g_4 - g_3) - (s_4 - s_3) \approx -10^{-3}$ ″/yr, which is related to Mars and the Earth, and $\sigma = (g_1 - g_5) - (s_1 - s_2) \approx -10^{-1}$ ″/yr, related to Mercury, Venus and Jupiter. They are the two most important small divisors appearing in the numerical integration as the frequency analysis of the solution shows since they appear with a large amplitude in the construction of quasiperiodic approximate interpolations of the orbits. The two corresponding arguments change several time from libration to circulation[7] over 200 Myrs, a phenomenon which is also typical of chaotic orbits (fig. 3).

[7] Near an approximate resonance one can introduce canonical coordinates such that the Hamiltonian is in its first approximation a pendulum Hamiltonian $H(p,q) = p^2 + \cos q$, $(p,q) \in \mathbb{R} \times \mathbb{T}^1$. Libration corresponds to homotopically trivial orbits ($H < 1$), circulation to homotopically non–trivial orbits ($H > 1$).

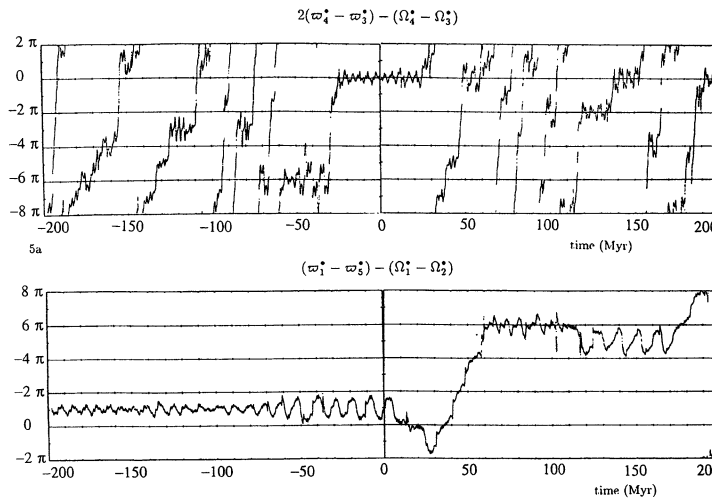


Fig. 3. Transition from libration to circulation of the argument corresponding to the secular resonances θ and σ

It should be stressed that the exponential divergence of the orbits results mostly from these repeated changes from libration to circulation of the resonant precession angles, which leads to a total indeterminacy of the orientation of the orbit in space. The eccentricities and inclinations variations are much much slower and become relevant only on a time scale of billions of years.

The numerical integrations later carried over by Laskar [L1994, L1995a, L1997] on time spans up to 25 Gyrs should be considered as an attempt to explore the chaotic zone where the Solar System evolves so as to have a qualitative description of the possible behaviour of the orbits on a time scale comparable to the age of the universe. The large planets (Jupiter, Saturn, Uranus and Neptun) have always very regular orbits whereas the eccentricities and inclinations of the inner planets show very large and irregular variations. In particular the eccentricity of Mercury reaches 0.5. Indeed exploiting the idea underlying the shadowing lemma [Y2] one can even find an orbit leading Mercury's orbit to intersect Venus' orbit in 3.5 Gyrs, thus leading to a collision or to escape.

2.4. The chaotic obliquity of the planets. Another kind of instabilities manifest themselves in the motion of the Solar System's planets. Because of their equatorial bulge the planets are subject to torques arising from the gravitational attraction of their satellites and of the Sun. This is at the origin of the precession of equinoxes (26000 years for the Earth). Moreover the obliquity of each planet is not fixed but is perturbed by the secular motion of the planet's orbit. Let $I_1 = I_2 < I_3$ be the

principal moments of inertia of the planet and assume that the axis of rotation of the planet is also its axis of maximum momentum of inertia I_3. The precession motion of the planet is given by $\frac{d\mathbf{H}}{dt} = \mathbf{L}$, where $\mathbf{H}$ is its spin angular momentum and $\mathbf{L} = \mathbf{L}(\lambda)$ is the torque exerted by the Sun. By averaging over the mean anomaly one obtains the secular equations of precession corresponding to the Hamiltonian[8]

$$\mathcal{H}(X, \psi, t) = \frac{\alpha}{2}(1 - e(t)^2)^{-3/2} X^2 + \sqrt{1 - X^2}(A(t) \sin \psi + B(t) \cos \psi)$$

where $\alpha = \frac{3m_0}{2a^3\nu} \frac{I_3 - I_1}{I_1}$, ν is the rotational angular velocity of the planet, m_0 is the solar mass. $X = \cos \varepsilon$, where ε is the obliquity, ψ is the precession angle. Here $A(t) + iB(t)$ is proportional to $\frac{d}{dt}(\xi_2(t) + i\eta_2(t))$ thus it depends on the change of orientation of the planet's orbital plane (inclination and longitude of the node).

When the effect of the perturbation to the orbital elements due to the other planets is not considered, the eccentricity e is constant and $A \equiv B \equiv 0$. The resulting Hamiltonian describes a free rotator, is clearly completely integrable and the obliquity is constant. X is an action variable and ψ is its canonically conjugated angle variable. If e is kept constant and $(A(t) \sin \psi + B(t) \cos \psi)$ is replaced with a single periodic term $A_0 \sin(\nu_0 t + \psi + \phi)$ then the resulting Hamiltonian is once again completely integrable. In the model studied by Laskar and Robutel [LR], $e(t)$ and $A(t) + iB(t)$ take into account the secular perturbations of the whole Solar System, modelled as a system with 15 degrees of freedom. Since the orbital solution [L1990] is not coupled with the precession variables it appears in the Hamiltonian $\mathcal{H}$ as an external time–dependent aperiodic forcing. The result is that all terrestrial planets could have experienced large, chaotic variations in obliquity at some time in the past. As frequency analysis shows, the obliquity of Mars is still in a large chaotic region, ranging from 0° to 60°, Mercury and Venus have been stabilized by tidial dissipation. In the case of Venus the crossing of a large chaotic zone extending from 0° to 90° during the slow–down process of his rotation speed due to dissipative effects may have contributed to the planet's retrograde rotation.

[8] In the case of the Earth the term $\frac{\alpha}{2}(1 - e(t)^2)^{-3/2} X^2$ must be changed because of the asymmetry of the earth ($I_1 < I_2$) and of the torque exerted by the Moon.

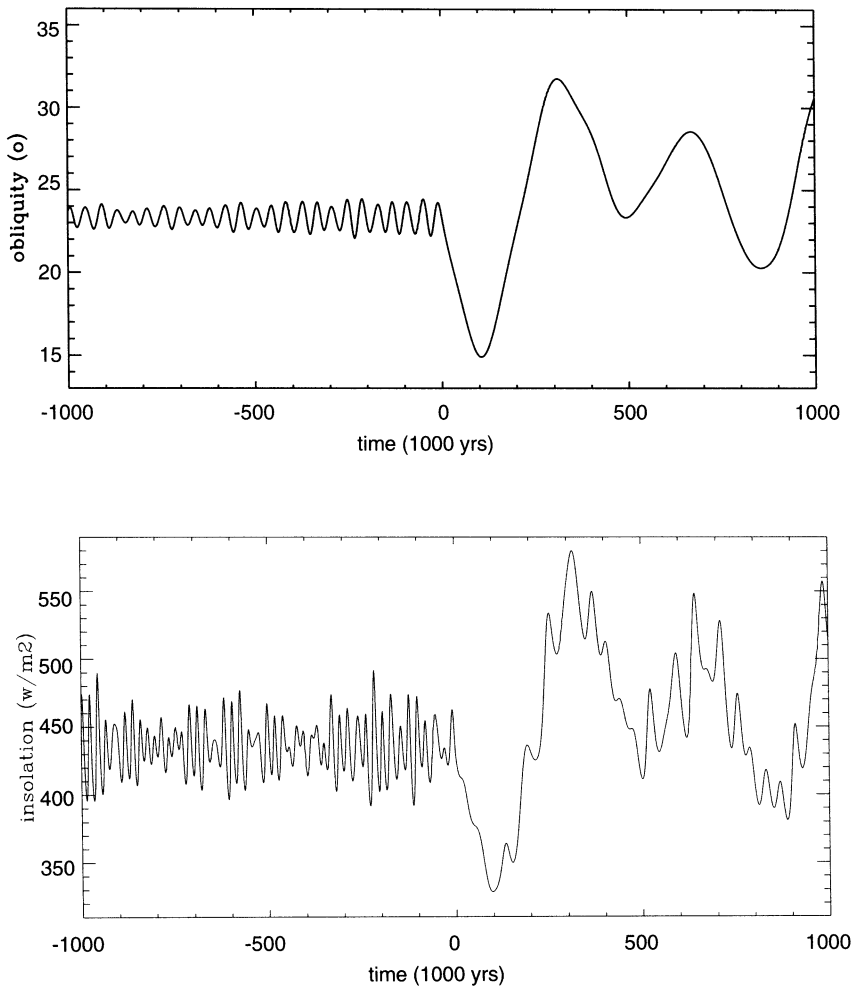


Fig. 4. Variation of the obliquity of the Earth (and of the insolation at northern latitudes) before and after "suppression" of the Moon

In its present state the Earth obliquity is essentially constant, with only small variations of $\pm 1.3°$ around the mean value of $23.3°$. Still, according to Milankovitch theory, this small variation is responsible for the ice ages. This relative stability is due to the fact that the Earth's precession frequency is presently far from being resonant with its secular frequencies, thus avoiding a large chaotic zone which extends from $60°$ to $90°$ in obliquity [LJR]. This zone would be even larger if the Moon were not present, extending from nearly $0°$ up to about $85°$ and leading to dramatic changes in climate (fig. 4).

2.5. Minor bodies in the Solar System. In addition to the 9 planets the Solar System is crowded with thousands of catalogued minor bodies (satellites, asteroids and comets). One can use simplified models, since their masses are negligible, but due to the variety of their canonical elements one needs to understand the dynamics in an even more global way.

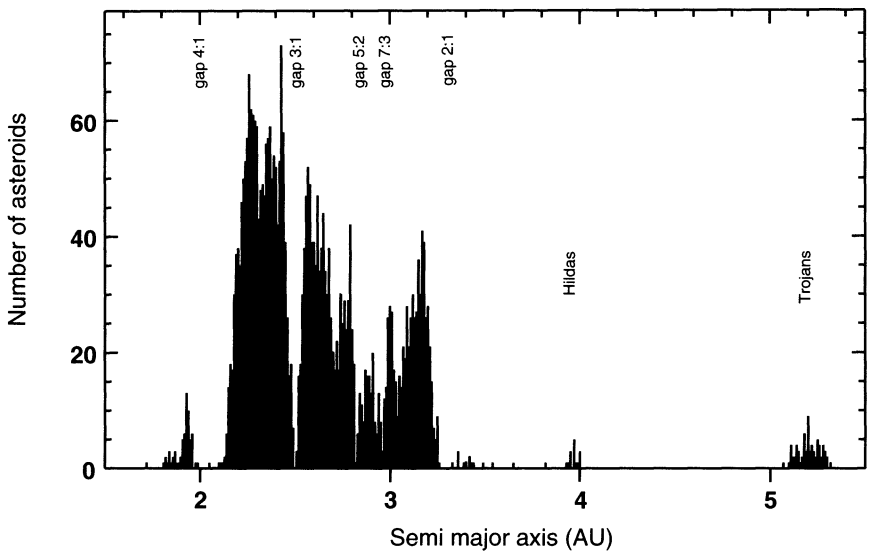


Fig. 5. Histogram of the number of asteroids against their semi–major axes

One of the first examples of chaotic behaviour was given by the chaotic tumbling of Hyperion, a small irregularly shaped satellite of Saturn whose strange rotational behaviour was detected during the encounter of the Voyager spacecraft with Saturn. In this case the dynamics is quite well understood and can be reduced to a pendulum Hamiltonian with a time–periodic perturbation [WPM]. For other satellites of the solar system one has a much more regular behaviour of their rotation since many KAM invariant tori populate the phase space [Ce].

The dynamics of the asteroids also raises a number of interesting problems. Their distribution, observed by plotting the number of asteroids against their semi–major axis length, shows gaps and accumulations first remarked by Kirkwood in 1867 (fig. 5). In this problem two and three–body mean motion resonances with Jupiter, Saturn or Mars play a very important role [Mo]. The chaotic diffusion in the inner asteroidal belt may even be at the origin of Mars and Earth–crossing asteroids$^{\heartsuit}$ [MN].

$^{\heartsuit}$ These are recently most popular among Hollywood's producers.

3. SOME OPEN PROBLEMS

The work of Laskar raises a number of challenging mathematical problems. What follows is a very rough list.

1. Can one give a rigorous justification of the frequency map analysis outside the set of KAM tori[9]? This implies to explore a largely unknown land ([He2], Section 6). At least for area–preserving maps the work of Mather [M] should provide some insight.

2. Prove a realistic stability result in the spirit of Nekhoroshev theorem for the system of the 4 outer planets and for a time comparable to the age of the Solar System.

3. Prove the existence of a KAM solution close to the Jupiter–Saturn system (see [Ro] for a discussion).

3'. Can one do the same for the 4 outer planets?

4. How long do the semi–major axes of the planets remain close to their initial values (see [Ni])?

5. Control the error term in the construction of the secular system numerically integrated by Laskar (is it possible to do it without answering to 4?)

6. Can one prove the existence of a transition chain which can possibly lead to the ejection of Mercury?

7. Can one find a mathematically satisfactory definition of stability for finite times, which reduces to orbital stability for infinite time, and which can be used to prove theorems? (An attempt in a somewhat well–defined context is given in [Gi]).

REFERENCES

[Ar1] V. I. ARNOL'D - *Small denominators and problems of stability of motion in classical celestial mechanics*, Russ. Math. Surv. **18** (1963), 85–193.

[Ar2] V. I. ARNOL'D - *Instability of dynamical systems with several degrees of freedom*, Sov. Math. Dokl. **5** (1964), 581–585.

[Ar3] V. I. ARNOL'D - *Huygens and Barrow, Newton and Hooke*, Birkhäuser (1990).

[Ar4] V. I. ARNOL'D (Ed.) - *Dynamical Systems III*, Springer–Verlag (1988).

[B] G. D. BIRKHOFF - *Dynamical Systems*, Amer. Math. Soc. Colloquium Pub. Vol. IX (1966).

[9] M. Herman remarked that the use of appropriate tools (e. g. spectral measures) from the spectral theory of dynamical systems (see [CFS], part III for an introduction) might be crucial.

[Bo] J.-B. BOST - *Tores invariants des systèmes dynamiques hamiltoniens*, Séminaire Bourbaki **639**, Astérisque **133–134** (1986), 113–157.

[Br] P. BRETAGNON - *Théorie du mouvement de l'ensemble des planètes. Solution VSOP82*, Astron. Astrophys. **114** (1982), 278–288

[Ch1] A. CHENCINER - *Intégration du problème de Kepler par la méthode de Hamilton–Jacobi: coordonnées "action–angles" de Delaunay*, Note S026, Bureau des Longitudes, Paris (1989).

[Ch2] A. CHENCINER - *À l'infini en temps fini*, Séminaire Bourbaki **832**, Astérisque **245** (1997), 323–353.

[Ce] A. CELLETTI - *Analysis of resonances in the spin–orbit problem in celestial mechanics: higher order resonances and some numerical experiments (Part II)* ZAMP **41** (1990), 453–479.

[CFS] I. P. CORNFELD, S. V. FOMIN, Ya. G. SINAI - *Ergodic Theory*, Grundlehren der mathematischen Wissenschaften **245**, Springer–Verlag, Berlin–Heidelberg–New York (1982).

[CMN] M. CARPINO, A. MILANI, A.M. NOBILI - *Long–term numerical integration and synthetic theories for the motion of outer planets*, Astron. Astrophys. **181** (1987), 182–194.

[Da] J. M. A. DANBY - *Fundamentals of Celestial Mechanics*, Willmann–Bell (1988).

[Gi] A. GIORGILLI - *On the problem of stability for near to integrable Hamiltonian systems*, ICM1998.

[GDFGS] A. GIORGILLI, A. DELSHAMS, E. FONTICH, L. GALGANI, C. SIMO - *Effective stability for a Hamiltonian system near an elliptic equilibrium point, with an application to the restricted three–body problem*, J. Diff. Eq. **77** (1989), 167–198.

[GS] A. GIORGILLI, CH. SKOKOS - *On the stability of the Trojan asteroids*, Astron. Astrophys. **317** (1997), 254–261.

[Ha] S.C. HARETU - *Sur l'invariabilité des grands axes des orbites planétaires*, Ann. Obs. Paris **XVIII** (1885), I1–I39.

[He1] M.R. HERMAN - *Sur les courbes invariantes par les difféomorphismes de l'anneau*, Astérisque **103–104** (1983).

[He2] M.R. HERMAN - *Some open problems in dynamical systems*, ICM1998.

[He3] M.R. HERMAN - *Stability in measure in the n–body problem*, seminar given at the ICTP Workshop on Dynamical Systems, Trieste (1998).

[K] A. N. KOLMOGOROV - *The General Theory of Dynamical Systems and Classical Mechanics*, address to the 1954 International Congress of Mathematicians, Amsterdam.

[Lg] J.L. LAGRANGE - *Sur l'altération des moyens mouvements des planètes*, Mem. Acad. Sci. Berlin 199 (1776) et Oeuvres Complètes **VI** Paris, Gauthier–Villars (1869), 255.

[Lp1] P.S. LAPLACE - *Mémoire sur les solutions particulières des équations différentielles et sur les inégalités séculaires des planètes*, (1772), Oeuvres complètes **IX** Paris, Gauthier–Villars (1895), 325.

[Lp2] P.S. LAPLACE - *Mémoire sur les inégalités séculaires des planètes et des satellites*, Mem. Acad. royale des Sciences de Paris (1784) and Oeuvres complètes **XI** Paris, Gauthier–Villars (1895), 49.

[L1989a] J. LASKAR - *Les variables de Poincaré et le développement de la fonction perturbatrice*, Note S026, Bureau des Longitudes, Paris (1989).

[L1989b] J. LASKAR - *A numerical experiment on the chaotic behaviour of the Solar System*, Nature **338** (1989), 237–238.

[L1990] J. LASKAR - *The chaotic behaviour of the Solar System: a numerical estimate of the size of the chaotic zones*, Icarus **88** (1990), 266–291.

[L1992] J. LASKAR - *A few points on the stability of the Solar System*, in "Chaos, Resonance and collective dynamical phenomena in the Solar System" Symposium IAU 152, S. Ferraz–Mello ed., Kluwer, Dordrechet (1992), 1–16.

[L1993] J. LASKAR - *Frequency analysis for multi–dimensional systems. Global dynamics and diffusion*, Physica **D 67** (1993), 257–281.

[L1994] J. LASKAR - *Large–scale chaos in the solar system*, Astron. Astrophys. **287** (1994), L9–L12.

[L1995a] J. LASKAR - *Large scale chaos and marginal stability of the Solar System*, in XI International Congress of Mathematical Physics, D. Iagolnitzer ed., International Press (1995), 75–120.

[L1995b] J. LASKAR - *Introduction to frequency map analysis*, Proceedings NATO Advanced Study Institute Hamiltonian systems with three or more degrees of freedom" S'Agaro, Spain, June 1995.

[L1996] J. LASKAR - *Chaos à grande échelle dans le Système Solaire et implications planétologiques*, C.R. Acad. Sci. Paris, **322** série IIa (1996), 163–180.

[L1997] J. LASKAR - *Large scale chaos and the spacing of inner planets*, Astron. Astrophys. **317** (1997), L75–L78.

[LFC] J. LASKAR, C. FROESCHLÉ, A. CELLETTI - *The measure of chaos by the numerical analysis of the fundamental frequencies. Application of the standard mapping*, Physica **D56** (1992), 253–269.

[LJB] J. LASKAR, F. JOUTEL, F. BOUDIN - *Orbital, precessional and insolation quantities for the Earth from -20Myr to +10 Myr*, Astron. Astrophys. **270** (1993), 522–533.

[LJR] J. LASKAR, F. JOUTEL, P. ROBUTEL - *Stabilization of the Earth's obliquity by the Moon*, Nature **361** (1993), 615–617.

[LRo] J. LASKAR, D. ROBIN - *Application of the frequency map analysis to the ALS*, Particle Accelerators **54** (1996), 183–192.

[LQT] J. LASKAR, T. QUINN, S. TREMAINE - *Confirmation of Resonant Structure in the Solar System*, Icarus **95** (1992), 148–152.

[LR] J. LASKAR, P. ROBUTEL - *The chaotic obliquity of the planets*, Nature **361** (1993), 608–612.

[Li1] J.E. LITTLEWOOD - *On the equilateral configuration in the restricted problem of three bodies*, Proc. London Math. Soc. **9** (1959), 343–372.

[Li2] J.E. LITTLEWOOD - *The Lagrange configuration in Celestial Mechanics*, Proc. London Math. Soc. **9** (1959), 525–543.

[Lo] P. LOCHAK - *Canonical perturbation theory via simultaneous approximations*, Russ. Math. Surv. **47** (1992), 57–133.

[M] J.N. MATHER - *Variational construction of connecting orbits*, Ann. Inst. Fourier **43** (1993), 1349–1386.

[MG] A. MORBIDELLI, A. GIORGILLI - *Superexponential Stability of KAM Tori*, J. Stat. Phys. **78** (1995), 1607–1617.

[MN] A. MORBIDELLI, D. NESVORNÝ - *Numerous weak resonances drive asteroids towards terrestrial planet orbits*, preprint (1998).

[Mo] M. MOONS - *Review of the dynamics in the Kirkwood gaps*, Cel. Mech. **65** (1997), 175–204.

[N] I. NEWTON - *Principia. Vol II. The System of the World*, Translation Motte, Univ. of California Press.

[Ne] N.N. NEKHOROSHEV - *An exponential estimate for the time of stability of nearly integrable Hamiltonian systems*, Russ. Math. Surveys **32** (1977), 1–65.

[Ni] L. NIEDERMAN - *Stability over exponentially long times in the planetary system*, Nonlinearity **9** (1996), 1703–1751.

[P] H. POINCARÉ - *Les Méthodes Nouvelles de la Mécanique Céleste*, tomes I–III, Paris Gauthier–Villars (1892, 1893, 1899).

[P2] H. POINCARÉ - *Leçons de Mécanique Céleste*, tome I, Paris Gauthier–Villars (1905).

[Po] S.D. POISSON - *Sur les inégalités séculaires des moyens mouvements des planètes*, Journal de L'Ecole Polytechnique **VIII** (1809), 1.

[Pol] H. POLLARD - *Celestial Mechanics*, The Carus Mathematical Monographs **18** (1976).

[Pö] J. PÖSCHEL - *Integrability of Hamiltonian systems on Cantor sets*, Comm. Pure Appl. Math. **35** (1982), 653–696.

[Ro] P. ROBUTEL - *Stability of the planetary three–body problem (part II)*, Celest. Mech. **62** (1995), 219–261.

[SM] C.L. SIEGEL, J. MOSER - *Lectures on Celestial Mechanics*, Springer–Verlag, Berlin–Heidelberg–New York (1971).

[St] S. STERNBERG - *Celestial Mechanics I*, W.A. Benjamin, New York (1969).

[SW] G.J. SUSSMAN, J. WISDOM - *Numerical evidence that the motion of Pluto is chaotic*, Science **241** (1988), 433–437.

[Wi] A. WINTNER - *The Analytical foundations of celestial mechanics*, Princeton University Press (1941).

[WPM] J. WISDOM, S.J. PEALE, F. MIGNARD - *The chaotic rotation of Hyperion*, Icarus **58** (1984), 137–152.

[Y1] J.–C. YOCCOZ - *An introduction to small divisors problems*, in "From Number Theory to Physics", M. Waldschmidt, P. Moussa, J.–M. Luck and C. Itzykson eds. Springer Verlag (1992), 659–679.

[Y2] J.–C. YOCCOZ - *Introduction to Hyperbolic Dynamics*, in "Real and Complex Dynamical Systems" B. Branner and P. Hjorth eds., NATO ASI Series **C464**, Kluwer, Dordrecht (1995), 265–291.

[Yo] L.S. YOUNG - *Ergodic Theory of Differentiable Dynamical Systems*, in "Real and Complex Dynamical Systems" B. Branner and P. Hjorth eds., NATO ASI Series **C464**, Kluwer, Dordrecht (1995), 293–336.

Stefano MARMI

Dipartimento di Matematica "U. Dini"
Università di Firenze
Viale Morgagni 67/A
50134 Firenze, Italy
marmi@udini.math.unifi.it

Astérisque

MICHEL BOILEAU

Uniformisation en dimension trois

Astérisque, tome 266 (2000), Séminaire Bourbaki,
exp. n° 855, p. 137-174

<http://www.numdam.org/item?id=SB_1998-1999__41__137_0>

© Société mathématique de France, 2000, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (http://smf4.emath.fr/Publications/Asterisque/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Séminaire BOURBAKI
51ème année, 1998-99, n° 855

Mars 1999

UNIFORMISATION EN DIMENSION TROIS

par **Michel BOILEAU**

INTRODUCTION

En 1977, W. Thurston a annoncé le théorème suivant :

Théorème de Thurston d'hyperbolisation.— *Soit M une variété compacte, orientable, de dimension trois. On suppose que toute sphère plongée S^2 borde une boule dans M et qu'il existe une surface orientable, de caractéristique d'Euler strictement négative, proprement plongée et dont le groupe fondamental s'injecte dans celui de M. L'intérieur de la variété M admet une structure hyperbolique complète si et seulement si tout sous-groupe abélien $\mathbb{Z}+\mathbb{Z}$ du groupe fondamental de M est conjugué à un sous-groupe du groupe fondamental d'une composante du bord de M.*

Ce théorème donne une condition simple sur la topologie d'une variété de dimension 3 pour que son intérieur admette une structure hyperbolique, c'est-à-dire une métrique riemannienne complète de courbure sectionnelle constante égale à -1. Il est au centre de beaucoup de résultats importants en dimension 3.

Un cas particulier important de ce théorème concerne une variété qui *fibre sur le cercle*, avec pour fibre une surface compacte orientable F et pour monodromie un difféomorphisme $\phi \in Diff^+(F)$ (cf. l'exposé de D. Sullivan [Su1]). La variété M ne dépend alors que de la classe de conjugaison de ϕ dans le groupe des difféotopies $\pi_0 Diff(F)$. La condition d'atoroïdalité est équivalente au fait que le difféomorphisme ϕ est *pseudo-Anosov* : son action sur l'ensemble des classes de conjugaison de lacets non triviaux dans $\pi_1(F)$ n'admet pas d'orbite périodique.

Le cas fibré peut alors s'énoncer sous la jolie forme suivante :

Théorème d'hyperbolisation pour les variétés fibrées sur le cercle ([Th4] ; voir aussi [Ota1]).— *Soit M une variété compacte, orientable, de dimension 3, qui admet une fibration sur le cercle de monodromie ϕ. L'intérieur de M porte une métrique hyperbolique complète de volume fini si et seulement si la monodromie ϕ est pseudo-Anosov.*

Historique de la preuve. Dans ces notes [Th0] Thurston a établi les techniques et les résultats sur lesquels se fondent la preuve de son théorème d'hyperbolisation. En 1980, il a présenté un plan détaillé et précis de celle-ci dans des notes sur les conférences qu'il avaient données à Bowdoin [Th1] (cf. [Th2], [Mor1]). Dans [Th3, Th4, Th5] il a écrit en détail la première partie de sa preuve (cf. § 4.1 le *théorème de l'image bornée*), ainsi que la preuve du cas fibré sur le cercle. Ces articles présentent les résultats nécessaires pour la preuve, ne figurant pas dans les notes [Th0]. Le reste de la preuve (cf. § 4.2) devait alors être reconstituée à l'aide de [Th1] et des chapitres 8, 9 et 13 de [Th0] (cf. [Mor1]).

En 1985, les travaux de F. Bonahon sur les bouts des variétés hyperboliques [Bon] (cf. § 4.2) ont permis de simplifier la partie de la preuve utilisant les chapitres 8 et 9 de Thurston.

À la même époque J. Morgan et P. Shalen [MS1, MS2, MS3] ont donné une preuve différente, de nature plus algébrique, des résultats principaux de [Th3, Th5] (cf. § 4.1).

En suivant une approche complètement différente de celle de Thurston, C. McMullen [McM1, McM2, McM3] a donné en 1989 une preuve du *théorème du point fixe* (cf. § 4), qui est le cœur de la démonstration du théorème d'hyperbolisation des variétés qui ne fibrent pas sur le cercle. Cette approche trouve son origine dans l'observation par J. Hubbard d'un lien entre la conjecture de Kra, sur la norme de l'opérateur Théta associé au revêtement d'une surface de Riemann, et le théorème du point fixe de Thurston.

En utilisant les travaux de McMullen, J.P. Otal [Ota2] a écrit en 1996 une preuve complète du théorème d'hyperbolisation des variétés qui ne fibrent pas sur le cercle. Entre-temps, il avait donné une nouvelle preuve pour le cas des variétés fibrées dans [Ota1].

À la même époque M. Kapovich [Ka] a écrit lui aussi une preuve complète du théorème d'hyperbolisation pour les variétés qui ne fibrent pas sur le cercle, mais en reconstituant l'approche originale de Thurston.

Dans ce texte M désignera toujours une variété compacte orientable. De même une surface sera par définition compacte et orientable.

Une variété M de dimension 3 est *irréductible* si tout plongement de la sphère $S^2 = \partial B^3 \hookrightarrow M$ se prolonge en un plongement de la boule B^3 dans M.

La variété M est *atoroïdale* si tout sous-groupe abélien de rang 2 du groupe fondamental de M est conjugué à un sous-groupe *périphérique* (i.e contenu dans le groupe fondamental d'une composante du bord).

Ces deux conditions, irréductibilité et atoroïdalité, sont nécessaires pour que l'intérieur de M admette une structure hyperbolique complète. C'est une conséquence directe de l'irréductibilité du revêtement universel $\mathbb{H}^3$ et de la classification des sous-groupes abéliens d'un sous-groupe discret et sans torsion de $PSL_2(\mathbb{C})$.

Une surface F, proprement plongée dans la variété irréductible M, est *essentielle* si le groupe fondamental de chaque composante connexe de F s'injecte dans celui de M (F est

dite alors *incompressible*) et qu'aucune composante connexe de F ne borde une boule ou ne peut être déformée dans le bord ∂M.

Une variété M est *hakenienne* si elle est compacte, irréductible et qu'elle contient une surface essentielle. Le théorème d'hyperbolisation est donc un théorème d'uniformisation des variétés hakeniennes atoroïdales, dont le groupe fondamental n'est pas virtuellement abélien.

Il existe beaucoup d'exemples importants de variétés hakeniennes : toute variété irréductible ayant un premier groupe d'homologie infini est hakenienne. En particulier toute variété irréductible à bord non vide est hakenienne. Un exemple typique est l'extérieur d'un nœud dans $\mathbb{S}^3$.

Le théorème d'hyperbolisation de Thurston s'applique à l'extérieur d'un nœud non trivial dans $\mathbb{S}^3$ qui n'est ni un nœud torique, ni un nœud satellite. C'est R. Riley [Ril], qui le premier a explicitement construit des exemples de structures hyperboliques sur le complémentaire de certains nœuds et entrelacs dans $\mathbb{S}^3$.

L'existence d'une structure hyperbolique complète de volume fini sur le complémentaire de ces nœuds est cruciale dans la preuve de la conjecture de Smith sur le redressement des actions non libres sur $\mathbb{S}^3$ d'un groupe cyclique fini. Pour la preuve de cette conjecture nous renvoyons aux actes édités par H. Bass et J. Morgan [BaM], ainsi qu'à l'exposé de J. Morgan [Mor2].

La condition d'être hakenienne n'est pas du tout nécessaire pour que l'intérieur de M porte une structure hyperbolique. C'est une conjecture toujours ouverte qu'elle peut être remplacée par la condition (cette fois-ci nécessaire) que le groupe fondamental de M est infini.

La variété hyperbolique fermée de Seifert et Weber [WS], obtenue en recollant par des isométries les faces d'un dodecaèdre hyperbolique d'angles dièdres $\frac{2\pi}{3}$, n'est pas hakenienne.

Le théorème de chirurgie hyperbolique de Thurston montre comment construire beaucoup de variétés fermées, hyperboliques qui ne sont pas hakeniennes. Partant d'une variété irréductible M, ne contenant aucune surface fermée essentielle et à bord un tore incompressible, Thurston déforme la structure hyperbolique complète (donnée par le théorème d'hyperbolisation) en une structure hyperbolique non complète, mais dont le complété est topologiquement une variété fermée, obtenue en collant un tore solide le long du bord ∂M.

Théorème de chirurgie hyperbolique ([Th0, Ch. 5]).— *Soit M une variété hakenienne atoroïdale et $T^2 \subset \partial M$ une composante torique incompressible. Excepté pour un nombre fini de recollements, une variété, obtenue en collant un tore solide le long du tore T^2, admet une structure hyperbolique complète.*

Pour plus de détails sur ce théorème nous renvoyons aux notes de Thurston [Th0, Ch. 5] et à l'exposé de M. Gromov [Gro].

L'importance des variétés atoroïdales dans l'étude des variétés de dimension 3 a été clairement mise en évidence par les travaux de Jaco-Shalen [JS] et Johannson [Jo]. Leurs résultats montrent que toute variété hakenienne M admet une décomposition (canonique), le long d'une famille finie (peut-être vide) de tores plongés, essentiels, disjoints et deux à deux non parallèles, en sous-variétés atoroïdales ou admettant une fibration en cercles (fibration de Seifert) ou en tores.

Ce résultat a donné le cadre topologique de la conjecture de géométrisation ci-dessous. Celle-ci est due à Thurston et domine depuis plus de 20 ans toute la topologie en dimension trois.

Les morceaux fibrés de la décomposition de Jaco-Shalen et Johannson peuvent être munis d'une métrique riemannienne homogène, complète, unique et localement isométrique à l'un des six modèles suivants : les géométries à courbure constante $\mathbb{S}^3$ et $\mathbb{E}^3$, la géométrie produit $\mathbb{H}^2 \times \mathbb{R}$, et enfin trois géométries fibrées modelées sur les groupes de Lie Nil, $\widetilde{SL_2(\mathbb{R})}$ et Sol.

Thurston [Th8] (voir aussi [Sco3]) a montré, qu'avec la géométrie hyperbolique, ce sont là les seules géométries homogènes que puisse admettre une variété irréductible et compacte de dimension 3. Ceci l'a conduit à formuler la conjecture suivante, qu'il a donc démontrée pour les variétés hakeniennes :

Conjecture de géométrisation (Thurston).— *L'intérieur d'une variété compacte, orientable, irréductible, de dimension 3, admet une décomposition canonique le long d'une famille finie de tores essentiels, en sous-variétés possédant une structure géométrique homogène complète.*

Cette conjecture englobe la conjecture de Poincaré, pour laquelle elle privilégie une approche beaucoup plus géométrique que topologique.

Sortant du cadre des variétés hakeniennes et des méthodes classiques de dimension 3, les travaux de P. Scott [Sco1, Sco2], P. Tukia [Tu], G. Mess [Me], D. Gabai [Ga1], A. Casson et D. Jungreis [CJ] ont permis de montrer qu'une variété irréductible, de groupe fondamental infini, admet une fibration de Seifert si et seulement si son groupe fondamental a un centre non trivial. Une conséquence de ce résultat est :

Théorème du tore ([CJ],[Ga1]).— *Soit M une variété de dimension 3, fermée, orientable et irréductible. Si la variété M a un groupe fondamental infini et qu'elle n'est pas atoroïdale, soit elle contient un tore plongé et essentiel, soit elle admet une fibration de Seifert.*

Ce résultat ramène la conjecture de géométrisation de Thurston à deux conjectures d'uniformisation, ne faisant plus intervenir que les géométries sphériques et hyperboliques. La première conjecture est une extension de la conjecture de Poincaré, englobant la conjecture de Smith sur le redressement des actions libres des sous-groupes finis de $Diff^+(\mathbb{S}^3)$.

Conjectures d'uniformisation.— *Soit M une variété orientable, fermée et irréductible :*

1. (Poincaré-Smith) *La variété M admet une métrique sphérique si et seulement si elle a un groupe fondamental fini.*

2. (Thurston) *La variété M admet une structure hyperbolique si et seulement si elle a un groupe fondamental infini et qu'elle est atoroïdale.*

À ce jour, ces conjectures restent largement ouvertes. Elles sont a priori indépendantes l'une de l'autre. C'est une idée fondamentale du programme de Thurston que d'avoir su lier ces deux conjectures en montrant que dans certains cas la géométrie sphérique peut être obtenue à partir d'un effondrement d'une suite de métriques hyperboliques singulières sur la variété considérée.

En 1981 Thurston [Th2] a annoncé le théorème suivant qui a pour conséquence immédiate la conjecture de Smith pour les actions cycliques finies, non libres sur $\mathbb{S}^3$.

Théorème de redressement des symétries ([Th6]).— *Soit M une variété de dimension 3 fermée, orientable, irréductible et atoroïdale. Soit $f \in Diff^+(M)$ un difféomorphisme périodique non trivial.. Si f n'agit pas librement sur M, la variété M admet une métrique sphérique ou hyperbolique invariante par f.*

Cet énoncé est un cas particulier du théorème des orbifolds de Thurston [Th6], qui montre l'analogue naturel de sa conjecture de géométrisation pour une orbifold compacte irréductible et dont le lieu de ramification est de dimension ≥ 1. Nous renvoyons à [BS] et [Th0, Ch. 13] pour la notion d'orbifold en général.

Dans le théorème ci-dessus, l'orbifold apparaît naturellement comme le quotient $\mathcal{O}$ de la variété M par l'action engendrée par le difféomorphisme f. Le point de départ de la preuve de ce théorème est là encore le théorème d'hyperbolisation qui permet de se ramener au cas où le complémentaire $\mathcal{O} - \Sigma$ de l'image Σ des points fixes de f admet une structure hyperbolique complète. On montre alors que soit cette structure hyperbolique peut être déformée en une structure hyperbolique dont le complété est topologiquement l'orbifold $\mathcal{O}$, soit l'orbifold $\mathcal{O}$ admet une structure sphérique. Pour plus de détails nous renvoyons à [BoP] et [CHK].

Nous achevons ce panorama de résultats sur l'uniformisation des variétés de dimension 3 par les résultats récents de D. Gabai [Ga2, Ga3], et D. Gabai, R. Meyerhoff, N. Thurston [GMT], qui eux aussi sortent du cadre des variétés hakeniennes. Une conséquence de leurs travaux est un théorème de rigidité hyperbolique, qui généralise en dimension 3 le théorème de rigidité de Mostow [Mos].

Théorème de rigidité hyperbolique ([GMT]).— *Une variété orientable fermée de dimension 3 admet une structure hyperbolique si et seulement si elle admet un revêtement fini qui a le type d'homotopie d'une variété hyperbolique.*

Voici un corollaire immédiat de ce théorème et du théorème d'hyperbolisation :

Corollaire.— *Si une variété orientable atoroïdale et fermée admet un revêtement fini qui est une variété hakenienne, elle admet une structure hyperbolique.*

La conjecture de Waldhausen suivante implique donc la conjecture de géométrisation de Thurston pour les variétés fermées, orientables et de groupe fondamental infini :

Conjecture du revêtement (Waldhausen).— *Toute variété fermée, orientable, irréductible et de groupe fondamental infini, admet un revêtement fini qui est une variété hakenienne.*

Cette conjecture est largement ouverte, même dans le cas d'une variété hyperbolique. On espère en fait qu'il existe toujours un revêtement fini ayant un premier groupe d'homologie infini.

Dans la preuve de beaucoup des résultats présentés dans cette introduction, le théorème d'hyperbolisation de Thurston joue un rôle central. On peut le considérer comme le résultat fondateur du programme de géométrisation des variétés de dimension 3, développé par Thurston. C'est pourquoi la suite de cet exposé lui est entièrement consacrée.

La preuve du théorème d'hyperbolisation, comme beaucoup de théorèmes profonds en dimension 3, utilise de façon essentielle l'hypothèse que la variété M est hakenienne. L'existence d'une surface essentielle F, proprement plongée dans M, permet de découper la variété le long de cette surface et d'obtenir ainsi une nouvelle variété M_F qui est irréductible. Si M_F n'est pas une réunion finie de boules, elle est hakenienne et le processus peut être itéré. Le théorème de finitude de Haken, qui borne le nombre de surfaces essentielles, disjointes, deux à deux non parallèles et proprement plongées dans la variété M, entraîne qu'en un nombre fini de pas ce processus donne une réunion finie de boules : une telle suite finie de variétés, obtenues par découpage, est appelée une *hiérarchie* de M. La preuve du théorème d'hyperbolisation s'effectue alors par récurrence sur la longueur maximale de certaines hiérarchies de M (cf. § 5). Le cœur de la démonstration est un théorème de recollement de structures hyperboliques (cf. § 3,4), auquel se réduit l'étape de récurrence par une jolie construction de ciselage et de miroitage de certaines composantes du bord (cf. § 5).

Ce texte ne prétend pas donner une preuve complète du théorème d'hyperbolisation de Thurston, mais seulement en présenter les étapes les plus marquantes.

À cet effet, les résultats ne sont pas énoncés dans leur plus grande généralité, mais uniquement dans le contexte de la preuve du théorème d'hyperbolisation.

La plupart des démonstrations sont seulement esquissées. Les détails complets peuvent être trouvés dans les articles originaux de W. Thurston [Th0, Th1, Th2, Th3, Th4, Th5], de C. McMullen [McM1, McM2], de J. Morgan et P. Shalen [MS1, MS2, MS3], ainsi que dans la monographie de M. Kapovich [Ka], l'article de J. Morgan [Mor1], et les monographies de J. P. Otal [Ota1, Ota2]. Il est clair que l'exposition présentée ici repose pour la plus grande partie sur ces références.

L'auteur tient à remercier L. Potyagailo pour l'avoir guidé d'un arbre à l'autre dans la théorie de Rips, et J.P. Otal pour son aide et ses suggestions qui ont permis d'améliorer substantiellement l'exposition. Il remercie aussi M. Burger et le Forschungsinstitut für Mathematik ETH Zürich pour leur hospitalité durant la préparation de ce texte.

1. VARIÉTÉS HYPERBOLIQUES ET GROUPES KLEINIENS

Définition.— *Un* groupe kleinien *est un sous-groupe discret de $PSL_2(\mathbb{C})$, de type fini et sans torsion.*

Nous supposerons toujours ici que le groupe kleinien Γ n'est pas virtuellement abélien.

L'action de Γ sur l'espace hyperbolique $\mathbb{H}^3$ s'étend en une action conforme sur la sphère à l'infini que l'on identifie à la sphère de Riemann $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$.

Le *domaine de discontinuité* Ω est l'ouvert Γ-invariant maximal de la sphère à l'infini, sur lequel l'action de Γ s'étend de façon libre et discontinue. Son complémentaire $\Lambda(\Gamma) = \hat{\mathbb{C}} - \Omega$ est l'adhérence de l'ensemble des points fixes des éléments de Γ. On l'appelle l'*ensemble limite de* Γ.

Le quotient $\overline{N} = \mathbb{H}^3 \cup \Omega/\Gamma$ est une *variété kleinienne*, dont l'intérieur $N = \mathbb{H}^3/\Gamma$ porte une structure hyperbolique complète et le bord $\partial N = \Omega/\Gamma$ une structure conforme. La variété $\overline{N}$ est naturellement orientée par l'orientation de $\mathbb{H}^3$.

Une surface de Riemann est de *type fini* si elle est conformément équivalente au complémentaire d'un nombre fini de points (appelés *pointes*) d'une surface de Riemann compacte. Elle est dite *hyperbolique* si chacune de ses composantes connexes a une caractéristique d'Euler strictement négative.

Le théorème suivant d'Ahlfors [Ah] (voir aussi [Su2]) est fondamental pour l'étude des variétés kleiniennes.

1.1. Théorème de finitude ([Ah])

Soit Γ un groupe kleinien non virtuellement abélien et dont le domaine de discontinuité Ω est non vide. La surface de Riemann Ω/Γ est hyperbolique de type fini. De plus, une pointe correspond à un élément parabolique de Γ. □

On dit que le groupe kleinien Γ admet un *parabolique accidentel* γ si $\gamma \in \Gamma$ est un élément parabolique qui correspond à un lacet de la surface Ω/Γ non homotope à une pointe.

Le *cœur convexe* $C(N)$ de la variété hyperbolique N est le quotient par Γ de l'enveloppe convexe $C(\Lambda)$ de Λ dans $\mathbb{H}^3$. C'est le plus petit convexe fermé, dont l'inclusion dans N est une homotopie d'équivalence. Ce n'est pas en général une sous-variété différentiable de N, car son bord peut être plissé le long de géodésiques. Il n'est pas non plus nécessairement de dimension 3. Par exemple il est de dimension 2 lorsque $\Gamma \subset PSL_2(\mathbb{R})$ est un groupe fuchsien, c'est-à-dire un sous-groupe discret et de covolume fini dans $PSL_2(\mathbb{R})$.

Pour éviter ce problème, on remplace souvent $C(N)$ par son δ-voisinage fermé $C_\delta(N)$, qui est l'ensemble des points de N à distance au plus δ de $C(N)$ pour un $\delta > 0$ fixé. Pour tout $\delta > 0$, $C_\delta(N)$ est une sous-variété de N, de dimension 3, de classe C^1, à bord strictement convexe et qui est difféomorphe à la variété kleinienne $\overline{N}$.

On note $N_{]0,\epsilon[}$ la partie ε-mince de N : c'est l'ensemble des points de N dont le rayon d'injectivité est $< \frac{\varepsilon}{2}$. Il existe alors une constante universelle ε_0, appelée constante de Margulis, telle que pour $\varepsilon \leq \varepsilon_0$ toute composante connexe de la partie ε-mince est soit un bout cuspidal de rang 1 ou 2, soit le voisinage tubulaire d'une géodésique (qu'on appelle *tube de Margulis*). On appelle alors *partie cuspidale* N^{cusp} de N, la réunion des bouts cuspidaux de la partie ε-mince pour $\varepsilon \leq \varepsilon_0$. On notera N_0 le complémentaire dans N de la partie cuspidale.

La variété hyperbolique N est *géométriquement finie* si son cœur convexe $C_\delta(N)$ est de volume fini. C'est équivalent à ce que la partie ε-épaisse $C_\delta(N) - N_{]o,\epsilon[}$ soit compacte. De plus, pour un ε suffisamment petit (dépendant de Γ), le type topologique de la variété compacte $M = C_\delta(N) - N_{]0,\epsilon[}$ est indépendant de ε et $P = \partial C_\delta(N) \cap N_{]0,\epsilon[}$ est une réunion finie d'anneaux et de tores correspondant respectivement aux cusps de rang 1 et 2.

Définition.— *On appelle* type topologique *de la variété hyperbolique géométriquement finie N la paire (M, P), où $P \subset \partial M$ est la sous-surface compacte de caractéristique d'Euler nulle, correspondant à la trace des cusps tronqués.*

La paire (M, P) a les propriétés topologiques suivantes :

1) M est une variété de dimension 3, compacte et irréductible ;

2) $P \subset \partial M$ est une sous-surface compacte, incompressible dans N, dont les composantes connexes sont des anneaux ou des tores. De plus P contient toutes les composantes toriques de ∂M ;

3) tout sous-groupe abélien $\mathbb{Z} + \mathbb{Z} \subset \pi_1(M)$ est conjugué au groupe fondamental d'une composante connexe torique de P ;

4) il n'existe pas d'anneau essentiel $(A, \partial A) \hookrightarrow (M, P)$, s'appuyant sur P ;

5) M n'est pas homéomorphe au produit $\mathbb{T}^2 \times [0, 1]$, excepté si $P = \mathbb{T}^2 \times \{0\}$, ni à un I-fibré tordu sur la bouteille de Klein.

Définition.— *Une paire (M, P) ayant les propriétés* 1) *à* 5) *est appelée* variété apprêtée atoroïdale. *La sous-surface $P \subset \partial M$ est appelée le* lieu parabolique *de la variété apprêtée et la surface compacte $\partial_0 M = \partial M - int(P)$ son* bord.

On dit que la variété apprêtée (M, P) est *acylindrique* si tout anneau essentiel dans $(M, \partial_0 M)$ peut être homotopé dans P.

Remarque : La variété apprêtée $(\mathbb{T}^2 \times [0, 1], \mathbb{T}^2 \times \{0\})$ est le type topologique d'un cusp de rang 2.

Définition.— *Une variété apprêtée (M, P) est dite* hyperbolique *si elle est le type topologique d'une variété hyperbolique géométriquement finie.*

Une version précise du théorème d'hyperbolisation de Thurston est :

1.2. Théorème d'hyperbolisation pour les variétés apprêtées ([Th1])

Soit M une variété hakenienne, compacte, orientable, de dimension 3. La variété apprêtée (M, P) est hyperbolique si et seulement si elle est atoroïdale.

Cette version du théorème d'hyperbolisation est nécessaire pour l'étape de récurrence. En effet les variétés apprêtées apparaissent naturellement au cours de la hiérarchie, par exemple lorsqu'on découpe une variété le long d'une surface essentielle s'appuyant sur une composante torique du bord.

La construction de miroitage décrite au § 5 permet à Thurston de ramener l'étape de récurrence à l'étape finale où le recollement fait intervenir tout le bord de la variété apprêtée.

2. LE THÉORÈME DE RECOLLEMENT

Dans toute la suite, (M, P) désigne une variété apprêtée atoroïdale à bord non vide et *strictement incompressible.* C'est-à-dire que $\partial_0 M$ est incompressible et qu'il n'existe pas d'anneau essentiel dans M s'appuyant à la fois sur $\partial_0 M$ et P.

On désignera toujours par τ une involution du bord $\partial_0 M$ de (M, P), qui échange les composantes connexes par paires et renverse leur orientation. C'est-à-dire que l'on a une partition $\partial_0 M = \partial_0^+ M \sqcup \partial_0^- M$ et un difféomorphisme renversant l'orientation $f : \partial_0^+ M \to \partial_0^- M$ tel que $\tau = (f, f^{-1})$.

On note alors M/τ la variété quotient, obtenue en identifiant deux à deux les composantes connexes de $\partial_0 M$ par τ. C'est une variété à bord vide ou une famille finie de tores $\partial(M/\tau) = P/\tau$. En tant que variété apprêtée, son bord est vide et le lieu parabolique est tout le bord $\partial(M/\tau)$.

Avec ces notations, l'étape principale de la preuve du théorème d'hyperbolisation de Thurston est le théorème de recollement suivant :

2.1. Théorème de recollement ([Th1])

Soit (M, P) une variété apprêtée, à bord non vide $\partial_0 M$ strictement incompressible. Soit $\tau : \partial_0 M \to \partial_0 M$ une involution qui échange les composantes connexes par paires en renversant leur orientation. On suppose que chaque composante connexe de la variété apprêtée (M, P) est hyperbolique. Alors, l'intérieur de la variété quotient M/τ admet une structure hyperbolique complète de volume fini si et seulement si elle est atoroïdale.

La preuve du théorème de recollement distingue deux cas suivant que la variété apprêtée (M, P) admet ou non un revêtement fini qui est le produit d'un intervalle par une surface compacte $(F, \partial F) \times I$. La démonstration dans les deux cas est entièrement différente.

Dans le premier cas, quitte à passer à un revêtement fini, on peut supposer que la variété apprêtée $(M, P)/\tau$ fibre sur le cercle, avec pour fibre la surface $(F, \partial F)$ et pour monodromie $f \in Diff^+(F)$, telle que $\tau = (f, f^{-1})$. La condition d'atoroïdalité est alors équivalente au fait que le difféomorphisme f est pseudo-Anosov.

Pour une preuve détaillée de ce cas nous renvoyons à [Ota1] et [Th2], ainsi qu'à [McM4] et [Su1].

Remarque : Dans beaucoup de cas la preuve du théorème d'hyperbolisation pour une variété non fibrée peut s'appliquer à un revêtement fini d'une variété fibrée. En fait, excepté pour une variété fermée dont le premier nombre de Betti est égal à 1, toute variété compacte qui fibre sur le cercle admet un revêtement fini qui contient une surface essentielle, proprement plongée et qui n'est la fibre d'aucune fibration sur le cercle. Pour un tel revêtement la preuve du cas non fibré s'applique. C'est une conjecture ouverte qu'un tel revêtement fini existe toujours.

À partir de maintenant, nous nous restreindrons au cas des variétés hakeniennes qui ne fibrent pas sur le cercle : c'est en quelque sorte le cas générique.

Nous allons expliquer maintenant comment Thurston ramène dans ce cas-là la preuve du théorème de recollement à un théorème de point fixe.

Rappelons que l'espace de Teichmüller $\mathcal{T}(F)$ d'une surface compacte connexe F est l'espace des classes d'équivalence de structures conformes d'aire finie sur int(F). Deux telles structures (F, s_1) et (F, s_2) sont équivalentes s'il existe une application conforme $h : F_1 \to F_2$ proprement homotope à l'identité.

On définit alors la distance de Teichmüller $\Delta(s_1, s_2)$ comme $\frac{1}{2} inf\{log(K(h))\}$, où $h : (F, s_1) \to (F, s_2)$ parcourt l'ensemble des applications quasi-conformes proprement homotopes à l'identité, et K(h) est l'excentricité de h.

Cette distance fait de $\mathcal{T}(F)$ un espace métrique complet, qui est homéomorphe à $\mathbb{R}^n$ pour $n = -3\chi(F)$, où $\chi(F)$ est la caractéristique d'Euler de F.

Dans toute la suite l'espace de Teichmüller $\mathcal{T}(\partial_0 M)$ désigne le produit des espaces de Teichmüller des composantes connexes de $\partial_0 M$. On le munit de la distance de Teichmüller définie comme le maximum des distances de Teichmüller pour les composantes connexes.

Voici l'analogue de l'espace de Teichmüller pour une variété apprêtée hyperbolique.

Définition.— *Soit (M, P) une variété apprêtée connexe, hyperbolique et à bord strictement incompressible. On note $\mathcal{GF}(M, P)$ l'ensemble* des classes d'équivalences de stuctures hyperboliques géométriquement finies sur (M, P). *C'est l'ensemble des paires $(N, [f])$, où N est une variété hyperbolique géométriquement finie et $[f]$ est la classe d'homotopie*

propre d'une équivalence d'homotopie de paires, respectant l'orientation, $f : (M, \partial M - P) \to (\overline{N}, \partial\overline{N})$. *Deux telles paires* $(N_1, [f_1])$ *et* $(N_2, [f_2])$ *sont équivalentes si la classe d'homotopie propre* $[(f_2)^{-1} \circ f_1]$ *peut être réalisée par une isométrie.*

Dans le cas général on définit $\mathcal{GF}(M, P)$ *comme le produit fini* $\prod \mathcal{GF}(M_i, P_i)$ *des ensembles* $\mathcal{GF}(M_i, P_i)$ *pour les composantes connexes* M_i *de* M.

Remarques : 1) Lorsque M est connexe, d'après A. Marden [Mard] et F. Waldhausen [Wa] la classe d'homotopie propre $[(f_2)^{-1} \circ f_1]$ peut toujours être réalisée par un homéomorphisme $\phi : \overline{N_1} \to \overline{N_2}$ dont la restriction au bord est quasi-conforme. L'ensemble $\mathcal{GF}(M, P)$ s'identifie donc à *l'espace des déformations quasi-conformes* d'une variété hyperbolique (ou d'un groupe kleinien) ayant pour type topologique la variété apprêtée (M, P).

2) Lorsque M est connexe, la remarque précédente permet de définir la distance de Teichmüller δ sur l'ensemble $\mathcal{GF}(M, P)$ comme suit :

$\delta((N_1, [f_1]), (N_2, [f_2])) = \frac{1}{2} inf\{log(K(\phi))\}$, où $\phi : N_1 \to N_2$ décrit tous les homéomorphismes dans la classe d'homotopie propre $[(f_2)^{-1} \circ f_1]$, qui sont quasi-conformes au bord, et $K(\phi)$ est l'excentricité de la restriction de ϕ au bord. Le théorème de Teichmüller montre que ce minimum est réalisé par un homéomorphisme, quasi-conforme au bord et d'excentricité minimale, unique dans la classe d'homotopie propre de $[(f_2)^{-1} \circ f_1]$.

Dans le cas général, on munit le produit $\mathcal{GF}(M, P)$ de la distance définie comme le maximum des distances sur les facteurs. Pour cette distance $\mathcal{GF}(M, P)$ est un espace métrique complet.

3) Si M est connexe, pour tout élément $(N, [f]) \in \mathcal{GF}(M, P)$ et pour toute composante connexe $P_i \subset P$, $f_*(\pi_1 P_i)$ est un sous-groupe parabolique du groupe kleinien $\pi_1(N)$, par définition de P. De plus toute transformation parabolique dans $\pi_1(N)$ est contenue, à conjugaison près, dans un des sous-groupes $f_*(\pi_1 P_i)$, où P_i parcourt l'ensemble des composantes connexes de P. Du fait que le bord $\partial_0 M$ de la variété apprêtée (M,P) est strictement incompressible, le groupe kleinien $\pi_1(N)$ n'admet pas de parabolique essentiel.

À partir de maintenant, et durant toute la preuve des théorèmes de recollement et du point fixe, nous supposerons que la variété apprêtée (M, P) *est connexe pour alléger les notations. Le cas où la variété apprêtée* (M, P) *n'est pas connexe ne présente pas vraiment de difficultés supplémentaires.*

Soit $\partial : \mathcal{GF}(M, P) \to \mathcal{T}(\partial_0 M)$ l'application de restriction qui à une variété hyperbolique géométriquement finie N, de type topologique (M, P), fait correspondre la structure conforme d'aire finie (d'après le théorème d'Ahlfors), induite sur le bord $\partial\overline{N}$ de la variété kleinienne $\overline{N}$ associée.

Le théorème suivant est crucial pour la preuve du théorème de recollement :

2.2. Théorème d'Ahlfors-Bers ([AB])

Soit (M, P) une variété apprêtée hyperbolique, à bord non vide strictement incompressible. Alors, l'application de restriction $\partial : \mathcal{GF}(M, P) \to \mathcal{T}(\partial_0 M)$ est un homéomorphisme. □

Remarque : Si $\partial_0 M$ est vide et que $\mathcal{GF}(M, P)$ n'est pas vide, il est réduit à un point d'après le théorème de rigidité de Mostow [Mos].

La preuve de la surjectivité de l'application de restriction ∂ découle du théorème d'existence et d'unicité des solutions de l'équation de Beltrami ([AB]). Son injectivité repose sur le théorème d'Ahlfors [Ah] qui montre que l'ensemble limite d'un groupe kleinien géométriquement fini et de covolume infini a une mesure de Lebesgue nulle.

Cette paramétrisation de l'espace $\mathcal{GF}(M, P)$ par l'espace de Teichmüller du bord $\mathcal{T}(\partial_0 M)$ a permis à Thurston de traduire le théorème de recollement en un théorème de point fixe pour un certain homéomorphisme de $\mathcal{T}(\partial_0 M)$ dans lui-même.

À cet effet, Thurston a introduit *l'application d'épluchage* $\sigma : \mathcal{T}(\partial_0 M) \to \mathcal{T}(\partial_0 \widetilde{M})$, où $\partial_0 \widetilde{M}$ est le bord de la variété apprêtée (M, P), avec l'orientation renversée. Nous définissons maintenant cette application. Dans ce qui suit l'indice F signifie la projection sur le facteur de $\mathcal{T}(\partial_0 M)$ correspondant à l'espace de Teichmüller $\mathcal{T}(F)$ de la composante $F \subset \partial_0 M$.

D'après le théorème d'Ahlfors-Bers, une classe de structures conformes $s \in \mathcal{T}(\partial_0 M)$ est induite par une structure hyperbolique M_s, géométriquement finie, sur (M, P). La variété hyperbolique M_s représente la classe $\partial^{-1}(s) = (M_s, [f_s]) \in \mathcal{GF}(M, P)$.

Comme M_s est sans parabolique accidentel, chaque composante connexe $F \subset \partial_0 M$ détermine un revêtement *quasi-fuchsien* M_s^F de la variété hyperbolique M_s, d'après B. Maskit [Mas1, Mas2]. C'est une variété hyperbolique géométriquement finie qui a pour type topologique la variété apprêtée produit $(F, \partial F) \times I$. En fait le groupe kleinien $\pi_1(M_s^F)$ est une déformation quasi-conforme d'un groupe fuchsien et son ensemble limite est une courbe de Jordan dans la sphère à l'infini $\hat{\mathbb{C}}$.

Soit $\tilde{f}_s : F \times I \to (\overline{M}_s^F, \partial \overline{M}_s^F)$ l'équivalence d'homotopie de paires obtenue par relèvement de $f_s : (M, \partial M - P) \to (\overline{M_s}, \partial \overline{M_s})$. L'application de restriction induit (d'après le théorème d'Ahlfors-Bers) un homéomorphisme $\partial_F : \mathcal{GF}((F, \partial F) \times I) \to \mathcal{T}(F) \times \mathcal{T}(\widetilde{F})$ tel que $\partial_F(M_s^F, [\tilde{f}_s]) = (s_F, s'_F)$, où $s_F \in \mathcal{T}(F)$ est la composante sur le facteur $\mathcal{T}(F)$ de $s \in \mathcal{T}(\partial_0 M)$.

Définition.— *L'application d'épluchage $\sigma : \mathcal{T}(\partial_0 M) \to \mathcal{T}(\partial_0 \widetilde{M})$ fait correspondre au point $s \in \mathcal{T}(\partial_0 M)$ le point $\sigma(s) \in \mathcal{T}(\partial_0 \widetilde{M})$ dont la composante sur le facteur $\mathcal{T}(\widetilde{F})$ est $\sigma(s)_F = s'_F$.*

Dans la suite on note $\tau^\star : \mathcal{T}(\partial_0 M) \to \mathcal{T}(\partial_0 \widetilde{M})$ l'involution induite par l'involution τ.

Thurston a traduit le critère de recollement de Maskit [Mas1, Mas2] en un critère de point fixe pour la composée $\tau^\star \circ \sigma : \mathcal{T}(\partial_0 M) \to \mathcal{T}(\partial_0 M)$.

2.3. Critère de recollement de Thurston ([Th1])

Soit (M,P) une variété apprêtée hyperbolique, vérifiant les hypothèses du théorème de recollement 2.1. *Si l'application $\tau^\star\circ\sigma : \mathcal{T}(\partial_0 M) \to \mathcal{T}(\partial_0 M)$ admet un point fixe, l'intérieur de la variété quotient M/τ admet une structure hyperbolique complète de volume fini.* □

Nous rappelons maintenant le critère de recollement de Maskit [Mas1, Mas2]. Nous nous restreignons au cas où la variété kleinienne $\overline{N}$ que l'on recolle est connexe. On note N son intérieur. On suppose que $\overline{N}$ est géométriquement finie, sans parabolique accidentel et à bord incompressible. On suppose aussi que $\overline{N}$ n'est pas un fibré en intervalles.

Soient F^+, F^- deux composantes connexes disjointes de $\partial\overline{N}$ et $f : F^+ \to F^-$ un difféomorphisme renversant les orientations. Le critère de Maskit donne une condition suffisante sur le recollement f pour que la variété quotient $\overline{N}/\tau$ soit une variété kleinienne géométriquement finie, où l'involution τ est l'involution $(f, f^{-1}) : (F^+ \sqcup F^-) \to (F^- \sqcup F^+)$.

Soient $\overline{N}^+$ et $\overline{N}^-$ les revêtements quasi-fuchsiens de $\overline{N}$, associés respectivement aux sous-groupes $\pi_1(F^+)$ et $\pi_1(F^-)$. L'hypothèse que $\overline{N}$ n'est pas un fibré en intervalles implique que ces revêtements sont de degré infini. On note $F_0^+ \subset \partial\overline{N}^+$ (respectivement $F_0^- \subset \partial\overline{N}^-$) la composante connexe dont la structure conforme s'identifie par revêtement à celle de $F^+ \subset \partial\overline{N}$ (respectivement $F^- \subset \partial\overline{N}$). On note $N^+ = int(\overline{N}^+)$ et $N^- = int(\overline{N}^-)$.

Avec les notations ci-dessus, le critère de Maskit s'énonce :

2.4. Critère de recollement de Maskit ([Mas1, Mas2])

Si l'équivalence d'homotopie propre $\tilde{f} : N^+ \to N^-$ induite par le difféomorphisme f est homotope à une isométrie $h : N^+ \to N^-$ dont l'extension au bord envoie F_0^+ sur $\partial\overline{N}^- - F_0^-$, la variété quotient $\overline{N}/\tau$ est une variété kleinienne géométriquement finie. □

Remarque : Soit (M,P) le type topologique de N. Par définition de l'application d'épluchage, la condition suffisante de Maskit est équivalente à la condition suivante :

$\tau^\star((\partial\rho_N)_{F^+}) = (\sigma(\partial\rho_N))_{F^+}$, où $\rho_N \in \mathcal{GF}(M,P)$ est la classe d'équivalence représentée par la variété hyperbolique $N = int(\overline{N})$.

Dans le théorème de recollement de toutes les composantes connexes de $\partial_0 M$ par paires, cette condition se traduit bien par le fait que $\partial\rho_N \in \mathcal{T}(\partial_0 M)$ est un point fixe de l'application $\tau^\star \circ \sigma$.

3. LE THÉORÈME DU POINT FIXE

Le critère de recollement de Thurston montre que le Théorème de recollement est une conséquence du théorème suivant :

3.1. Théorème de Thurston du point fixe

Soit (M,P) une variété apprêtée hyperbolique, qui n'est pas fibrée en intervalles et dont le bord $\partial_0 M$ est non vide et strictement incompressible. Soit $\tau : \partial_0 M \to \partial_0 M$ une

involution qui échange les composantes par paires en renversant leur orientation. Alors, l'application $\tau^\star \circ \sigma : \mathcal{T}(\partial_0 M) \to \mathcal{T}(\partial_0 M)$ *a un point fixe si et seulement si la variété quotient* M/τ *est atoroïdale.*

La proposition suivante montre que l'application $\tau^\star \circ \sigma$ contracte strictement la distance de Teichmüller sur l'espace métrique complet $\mathcal{T}(\partial_0 M)$. C'est une conséquence du fait que l'application quasi-conforme de Teichmüller, d'excentricité minimale dans sa classe d'homotopie propre, est analytique réelle excepté en un nombre fini de points.

3.2. Proposition.— *Soit* (M, P) *une variété apprêtée hyperbolique, vérifiant les hypothèses du théorème du point fixe. Alors, l'application* $\tau^\star \circ \sigma : \mathcal{T}(\partial_0 M) \to \mathcal{T}(\partial_0 M)$ *contracte strictement la distance de Teichmüller sur* $\mathcal{T}(\partial_0 M)$:

$\Delta(\tau^\star \circ \sigma(s), \tau^\star \circ \sigma(s')) < \Delta(s, s')$. □

Remarques : 1) En général la contraction n'est pas uniforme sur $\mathcal{T}(\partial_0 M)$.

2) Comme $\tau^\star : \mathcal{T}(\partial_0 M) \to \mathcal{T}(\partial_0 \widetilde{M})$ est une isométrie pour la distance de Teichmüller, il suffit de montrer que l'application d'épluchage σ contracte la distance de Teichmüller.

3) Si la variété apprêtée (M, P) est un fibré en intervalles sur une surface compacte orientable $(F, \partial F)$, l'application d'épluchage $\sigma : \mathcal{T}(F) \times \mathcal{T}(\widetilde{F}) \to \mathcal{T}(\widetilde{F}) \times \mathcal{T}(F)$ est définie par $\sigma(s, \tilde{s}') = (\tilde{s}', s)$. C'est une isométrie pour la distance de Teichmüller.

La proposition 3.2 implique l'unicité du point fixe pour l'application $\tau^\star \circ \sigma$ s'il existe. De plus, elle ramène son existence à celle d'un sous-ensemble compact de $\mathcal{T}(\partial_0 M)$ globalement invariant par $\tau^\star \circ \sigma$.

Pour un point $s \in \mathcal{T}(\partial_0 M)$ on note $\mathcal{O}(s) = \{(\tau^\star \circ \sigma)^n(s); n \in \mathbb{Z}\}$ son orbite par l'action de $\tau^\star \circ \sigma$. Son adhérence $\overline{\mathcal{O}}(s)$ est un fermé invariant par $\tau^\star \circ \sigma$. C'est donc un candidat naturel pour être un compact invariant par $\tau^\star \circ \sigma$. Comme l'espace métrique $\mathcal{T}(\partial_0 M)$ est complet, il suffit de trouver une orbite $\mathcal{O}(s)$ bornée. Le théorème suivant est le cœur de la preuve du théorème du point fixe.

3.3. Théorème de l'orbite bornée

Soit (M, P) *une variété apprêtée hyperbolique, vérifiant les hypothèses du théorème du point fixe. Si la variété quotient* M/τ *est atoroïdale, l'orbite* $\mathcal{O}(s)$ *de tout point* $s \in \mathcal{T}(\partial_0 M)$ *est bornée dans* $\mathcal{T}(\partial_0 M)$ *pour la distance de Teichmüller.*

Il existe deux preuves très différentes de ce théorème.

La première preuve est due à Thurston et utilise beaucoup d'outils géométriques qu'il a forgés à cet effet et qui sont maintenant fondamentaux dans l'étude des variétés hyperboliques de dimension 3. C'est en particulier le cas des *surfaces plissées* et des *laminations mesurées terminales*, pour décrire les bouts des variétés hyperboliques qui ne sont pas géométriquement finies. On peut pour ces notions se reporter aux chapitres 8 et 9 de [Th0], à [CEG] et à [Bon].

Une autre preuve a été donnée en 1989 par C. McMullen. Elle suit une approche complètement différente, plus directe et plus proche de la théorie de Teichmüller classique. C. McMullen a d'abord prouvé une version plus générale de la conjecture de Kra, en montrant que la norme de l'opérateur Théta, associé à un revêtement de surfaces de Riemann hyperboliques, est < 1 si et seulement si le revêtement n'est pas moyennable [McM1]. À partir de ce résultat, il obtient une très jolie majoration de la norme de la dérivée de l'application d'épluchage en un point $s \in \mathcal{T}(\partial_0 M)$: $\|d\sigma(s)\| < 1 - \eta$, où la constante η ne dépend que de la caractéristique d'Euler de $\partial_0 M$ et de la systole de la structure conforme s sur $\partial_0 M$. L'atoroïdalité de la variété M/τ lui permet alors de conclure par un argument analogue à celui du lemme 4.2.7 (cf. [McM2]; voir aussi [Ota2, § 6]).

Voici un énoncé du théorème de McMullen qui implique le théorème de l'orbite bornée :

3.4. Théorème de McMullen de contraction ([McM2])

Soit (M, P) une variété apprêtée hyperbolique, vérifiant les hypothèses du théorème du point fixe. On suppose que la variété quotient M/τ est atoroïdale. Il existe un entier $K > 0$ (ne dépendant que de la topologie de $\partial_0 M$) tel que $\|d(\tau^\star \circ \sigma)^K(s)\| \leq c < 1$ en tout point $s \in \mathcal{T}(\partial_0 M)$ pour lequel $\Delta(s, \tau^\star \circ \sigma(s)) \leq \Delta(s_0, \tau^\star \circ \sigma(s_0))$, où la constante c est uniforme pour tous ces points et $s_0 \in \mathcal{T}(\partial_0 M)$ est un point arbitrairement choisi.

Remarque : Dans le cas où la variété apprêtée (M, P) est acylindrique, McMullen [McM2] montre le résultat plus fort que l'application $\tau^\star \circ \sigma$ est une contraction de constante uniforme sur tout l'espace de Teichmüller $\mathcal{T}(\partial_0 M)$. Il faut remarquer que dans ce cas-là l'atoroïdalité de la variété quotient M/τ est toujours vérifiée indépendamment de l'involution τ.

Pour une preuve du Théorème de contraction, nous renvoyons aux articles originaux de McMullen [McM1, McM2], ainsi qu'à la monographie de J.P. Otal [Ota2]. Pour une présentation concise des résultats de McMullen le lecteur pourra se reporter à [McM3].

Nous allons maintenant expliquer la preuve du théorème de l'orbite bornée suivant l'approche de Thurston. Nous suivrons pour cela la présentation qui en est donnée par M. Kapovich dans son livre [Ka].

4. LA PREUVE DU THÉORÈME DE L'ORBITE BORNÉE

Soit (M, P) une variété aprêtée hyperbolique.

Définition.— *$\mathcal{X}(M, P)$ est l'espace des classes de conjugaison de représentations fidèles et discrètes de $\pi_1(M)$ dans $PSL_2(\mathbb{C})$ telles que $\pi_1(P_i)$ soit parabolique pour toute composante connexe $P_i \subset P$.*

On munit cet espace de la topologie quotient de la topologie de la convergence simple sur l'ensemble des morphismes du groupe discret $\pi_1(M)$ dans $PSL_2(\mathbb{C})$.

Comme $\pi_1(M)$ n'est pas virtuellement abélien, $\mathcal{X}(M,P)$ est un espace métrisable et complet. C'est en fait un sous-espace fermé de la variété affine des caractères de $\pi_1(M)$ dans $PSL_2(\mathbb{C})$.

En oubliant la condition d'être géométriquement finie, on obtient une application naturelle de l'espace $\mathcal{GF}(M,P)$ sur l'ouvert suivant de $\mathcal{X}(M,P)$: l'ensemble des classes de conjugaison de représentations de $\pi_1(M)$ dans $PSL_2(\mathbb{C})$ provenant des déformations quasi-conformes d'une structure hyperbolique géométriquement finie de type topologique (M,P). Lorsqu'on munit $\mathcal{GF}(M,P)$ de la distance de Teichmüller δ cette application est un homéomorphisme.

Ceci permet d'identifier $\mathcal{GF}(M,P)$ à un sous-ensemble ouvert de $\mathcal{X}(M,P)$. Cependant L. Bers [Ber] a montré que $\mathcal{GF}(M,P)$ n'est pas complet pour la topologie induite.

La preuve du théorème de l'orbite bornée, suivant Thurston, consiste à montrer que, pour une orbite $\mathcal{O}(s) \subset \mathcal{T}(\partial_0 M)$, son image $\partial^{-1}(\mathcal{O}(s))$ est relativement compacte dans $\mathcal{GF}(M,P)$. Elle comporte deux étapes :

1) On montre d'abord que l'image $\partial^{-1}(\mathcal{O}(s))$ est bornée, donc relativement compacte, dans l'espace $\mathcal{X}(M,P)$.

2) Puis on montre que l'adhérence de $\partial^{-1}(\mathcal{O}(s))$ dans $\mathcal{X}(M,P)$ est en fait contenue dans $\mathcal{GF}(M,P)$.

Aucune de ces deux étapes ne s'obtient aisément. Elles correspondent aux deux énoncés suivants :

4.1. Théorème de l'image bornée

Soit (M,P) une variété apprêtée hyperbolique, vérifiant les hypothèses du théorème du point fixe. Si la variété quotient M/τ est atoroïdale, l'image $\partial^{-1}(\mathcal{O}(s))$ de l'orbite d'un point $s \in \mathcal{T}(\partial_0 M)$ est bornée dans $\mathcal{X}(M,P)$.

4.2. Théorème de la limite géométrique

Soit (M,P) une variété apprêtée hyperbolique, vérifiant les hypothèses du théorème du point fixe. Si la variété quotient M/τ est atoroïdale, les deux assertions suivantes sont vérifiées :

1) *Tout point d'accumulation $\rho_\infty \in \overline{\partial^{-1}(\mathcal{O}(s))} \subset \mathcal{X}(M,P)$, est limite d'une suite $\{\rho_k\}_{k\in\mathbb{N}} \subset \partial^{-1}(\mathcal{O}(s))$ telle que la suite de variétés hyperboliques pointées $\{(\mathbb{H}^3/\rho_k(\pi_1 M), x_k)\}_{k\in\mathbb{N}}$ converge géométriquement vers la variété hyperbolique pointée $(\mathbb{H}^3/\rho_\infty(\pi_1 M), x_\infty)$. Les points bases sont choisis avec un rayon d'injectivité minoré par la constante de Margulis.*

2) *Tout point d'accumulation $\rho_\infty \in \overline{\partial^{-1}(\mathcal{O}(s))} \subset \mathcal{X}(M,P)$ appartient à $\mathcal{GF}(M,P)$.*

Le terme de *convergence géométrique* fait référence à la convergence bilipschitz pointée, introduite par M. Gromov [GLP] sur l'espace des variétés riemanniennes de dimension donnée et à courbure sectionnelle pincée.

Nous allons maintenant décrire la logique de la preuve du théorème de l'image bornée.

4.1. Preuve du théorème de l'image bornée

Nous suivons la preuve donnée par M. Kapovich dans son livre ([Ka]), via la théorie des actions de groupes sur les arbres réels et la théorie de Rips. Il s'agit en quelque sorte d'une approche duale de celle de Thurston dans [Th3, Th5], qui utilise des triangulations du revêtement universel de M, équivariantes par l'action de $\pi_1(M)$.

J. Morgan et P. Shalen dans [MS1, MS2, MS3] sont les premiers à avoir développé et utilisé la théorie des actions de groupes sur les arbres réels pour compactifier l'espace $\mathcal{X}(M,P)$. Ils ont pour cela développé tout un programme d'étude des actions de groupes, sans point fixe global et à petits stabilisateurs d'arêtes, sur les arbres réels, qui est à l'origine du théorème de Rips.

Nous rappelons maintenant quelques définitions sur les actions de groupes sur les arbres. Pour plus de détails nous renvoyons à l'exposé de F. Paulin [Pau2].

Un arbre réel est un espace métrique T tel que toute paire de points distincts $(x,y) \in T \times T$ soit les extrémités d'un unique arc topologique dans T et que cet arc soit isométrique à un intervalle de $\mathbb{R}$. *Un arbre simplicial* est un 1-complexe simplicial, connexe et simplement connexe, muni de la métrique des chemins, rendant chaque arête isométrique à un intervalle.

Une action isométrique d'un groupe G sur un arbre réel est dite *triviale* si elle admet un point fixe global. Elle est dite *minimale* si elle ne laisse invariant aucun sous-arbre propre. Elle est dite *petite* si le stabilisateur de toute arête (non réduite à un point) ne contient pas de sous-groupe libre non abélien.

Pour tout élément $g \in G$ on définit la *longueur de translation* $\ell_T(g) = \inf\{d(x,g(x)), x \in T\}$. Cette longueur est atteinte dans T. L'élément g est dit *elliptique* si $\ell_T(g) = 0$.

Soit G un groupe de type fini, non virtuellement nilpotent. Les actions isométriques de G sur des arbres réels apparaissent naturellement pour compactifier l'espace $\mathcal{X}(G)$ des classes de conjugaison fidèles et discrètes de G dans $PSL_2(\mathbb{C})$. La première construction est due à M. Culler et P. Shalen [CS], en utilisant les points idéaux des courbes de la variété des caractères de G dans $PSL_2(\mathbb{C})$. Elle a été ensuite utilisée et généralisée par J. Morgan et P. Shalen [MS1, MS2, MS3].

Nous allons ici rappeler rapidement la méthode, plus géométrique, due indépendamment à M. Bestvina [Bes] et F. Paulin [Pau1]. Pour plus de détails nous renvoyons à [Pau2].

Étant donné un groupe G de type fini, on se donne un système fini $\mathcal{S}$ de générateurs. La construction qui suit est indépendante du système de générateurs choisi. Pour une suite de représentations $\{\rho_n\}_{n\in\mathbb{N}} \subset \mathcal{X}(G)$, on définit la suite de nombres réels $\lambda_n = inf\{max\{d(x,\rho_n(\gamma)(x)), \gamma \in \mathcal{S}\}; x \in \mathbb{H}^3\}$. Comme G n'est pas virtuellement nilpotent, il existe un point $x_n \in \mathbb{H}^3$ tel que $\lambda_n = max\{d(x_n,\rho_n(\gamma)(x_n)), \gamma \in \mathcal{S}\}$.

Si la suite de représentations $\{\rho_n\}_{n\in\mathbb{N}}$ n'est pas bornée dans $\mathcal{X}(G)$, $\lim_{n\to\infty}\lambda_n = +\infty$. On considère la suite d'espaces métriques pointés réechelonnés $(\frac{1}{\lambda_n}\mathbb{H}^3, x_n)_{n\in\mathbb{N}}$, munis pour chaque n de l'action isométrique de G induite par la représentation ρ_n. Il existe alors

une suite extraite, convergeant pour la topologie de Hausdorff-Gromov équivariante vers un arbre réel T (non réduit à un point) et muni d'une action isométrique de G qui est minimale, petite et non triviale. De plus : $\lim_{n\to\infty} \ell(\rho_n(g)) = \ell_T(g), \forall g \in G$, où $\ell(\rho_n(g))$ est la longueur de translation de $\rho_n(g) \in PSL_2(\mathbb{C})$.

On dit que la suite (non bornée) de représentations $\{\rho_n\}_{n\in\mathbb{N}} \subset \mathcal{X}(G)$ admet une suite extraite qui converge (pour la topologie de Hausdorff-Gromov équivariante) vers une action (isométrique) de G sur un arbre réel.

La proposition suivante découle de la construction précédente :

4.1.1. Proposition

Soit (M, P) une variété apprêtée hyperbolique, vérifiant les hypothèses du théorème du point fixe. Si l'image $\partial^{-1}(\mathcal{O}(s))$ de l'orbite d'un point $s \in \mathcal{T}(\partial_0 M)$ n'est pas bornée dans $\mathcal{X}(M, P)$, il existe une suite de représentations $\{\rho_n\} \subset \partial^{-1}(\mathcal{O}(s))$ qui converge pour la topologie Hausdorff-Gromov équivariante vers une action isométrique minimale, petite et non triviale, de $\pi_1(M)$ sur un arbre réel T. De plus, pour chaque composante connexe $P_i \subset P$ l'action du sous-groupe $\pi_1(P_i)$ est triviale sur T. □

La démonstration du théorème de l'image bornée résulte alors, par contradiction, de la proposition suivante :

4.1.2. Proposition

Soit (M, P) une variété apprêtée hyperbolique, vérifiant les hypothèses du théorème du point fixe. On suppose que la variété quotient M/τ est atoroïdale. Alors, toute action isométrique de $\pi_1(M)$ sur un arbre réel T, obtenue comme limite (pour la topologie Hausdorff-Gromov équivariante) d'une suite de représentations dans $\partial^{-1}(\mathcal{O}(s))$, $s \in \mathcal{T}(\partial_0 M)$, est triviale.

Démonstration de la Proposition 4.1.2 (cf. [Ka, §, 14]) : On considère une action isométrique de $\pi_1(M)$ sur un arbre réel T, vérifiant les hypothèses de la Proposition 4.1.2. La démonstration consiste à montrer d'abord que tout sous-groupe périphérique (i.e. provenant du bord) agit trivialement sur l'arbre réel T. Puis on utilise la décomposition de Jaco-Shalen [JS] et Johannson [Jo] de la variété apprêtée (M, P), pour conclure que l'action du groupe $\pi_1(M)$ est triviale sur T.

Étant donnée une variété apprêtée atoroïdale (M, P), à bord strictement incompressible, une conséquence de la théorie de Jaco-Shalen [JS] et Johannson [Jo] est l'existence d'une paire de sous-variétés $(W, \partial_0 W) \subset (M, \partial_0 M)$ ayant les propriétes suivantes :

1) Chaque composante connexe de $(W, \partial_0 W)$ est un fibré en intervalles, qu'on appellera *une fenêtre* de la variété apprêtée (M, P).

2) La frontière $\partial_1 W = \overline{\partial W - (\partial_0 W \cup P)}$ est une réunion finie d'anneaux essentiels proprement plongés dans $(M, \partial_0 M)$.

3) Chaque composante connexe M_i de $\overline{M-W}$ est soit un tore solide V dont le bord contient un anneau essentiel de $\partial_1 W$, soit *acylindrique* (tout anneau essentiel, proprement plongé dans $(M_i, \overline{\partial M_i \cap \partial_0 M})$ est parallèle dans M_i à un des anneaux $(\partial_1 W \cap M_i, \partial(\partial_1 W \cap M_i))$.

4) La paire $(W, \partial_0 W)$ est unique, à isotopie de paire près, dans $(M, \partial_0 M)$.

Remarque : La sous-variété caractéristique de Jaco-Shalen et Johannson pour la paire $(M, \partial_0 M)$ est la réunion de W et des composantes connexes de $\overline{M-W}$ qui sont des tores solides. De plus, la paire $(M, \partial_0 M)$ est acylindrique si et seulement si W est vide.

Le théorème suivant de J. Morgan et P. Shalen est le cœur de la démonstration de la Proposition 4.1.2. Il donne une nouvelle démonstration des résultats principaux de [Th3, Th5].

4.1.3. Théorème ([MS3])

Soit (M, P) une variété apprêtée hyperbolique, à bord strictement incompressible et qui n'est pas un fibré en intervalles. Soit $W \subset M$ la réunion des fenêtres de (M, P). On suppose que le groupe fondamental $\pi_1(M)$ admet une action isométrique minimale et petite sur un arbre réel T, telle que l'action induite par le groupe fondamental de chaque composante connexe de P est triviale. Alors l'action induite par le groupe fondamental de chaque composante connexe de $\overline{M-W}$ est triviale sur T.

Le résultat de compacité suivant, dû à Thurston, est un corollaire direct de la proposition 4.1.1 et du théorème de Morgan et Shalen.

4.1.4. Corollaire ([Th3], [MS3])

Soit (M, P) une variété apprêtée hyperbolique, à bord strictement incompressible. Si (M, P) est acylindrique, l'espace $\mathcal{X}(M, P)$ est compact. □

Démonstration du Théoreme 4.1.3.— Nous allons maintenant esquisser la preuve de ce théorème suivant un argument qui nous a été expliqué par L. Potyagailo. Cet argument utilise une version plus précise du théorème de Rips, qui est due à V. Guirardel [Gui].

Le théorème de Rips fournit une action minimale, non triviale et petite de $\pi_1(M)$ sur un arbre simplicial (cf. [BeF], [Pau2]). Le théorème de Guirardel précise ce résultat, en fournissant des actions de $\pi_1(M)$ sur des arbres simpliciaux qui sont du même type et arbitrairement proches pour la topologie de Hausdorff-Gromov équivariante de l'action de $\pi_1(M)$ sur l'arbre réel T.

La preuve du théorème 4.1.3 découle alors du fait que les actions non triviales de $\pi_1(M)$ sur ces arbres simpliciaux induisent des scindements non triviaux du groupe $\pi_1(M)$ le long de sous-groupes $\mathbb{Z}$ correspondant à des anneaux essentiels de $\partial_1 W$. La maximalité de la sous-variété W impose alors que les groupes fondamentaux $\pi_1(M_i)$ des composantes connexes de $\overline{M-W}$ agissent trivialement sur ces arbres simpliciaux.

La continuité de l'application $T \to \ell_T(\gamma)$ dans la topologie Hausdorff-Gromov équivariante, pour un élément fixé $\gamma \in \pi_1(M)$, et le théorème de Guirardel [Gui] montrent alors que les actions des groupes $\pi_1(M_i)$ sur l'arbre réel T doivent être triviales. □

Une conséquence de l'atoroïdalité de la variété quotient M/τ et du théorème 4.1.3 de Morgan et Shalen est :

4.1.5. Corollaire ([Ka, §14.17])

Soit (M, P) une variété apprêtée hyperbolique, vérifiant les hypothèses de la Proposition 4.1.2. *L'action du groupe fondamental de toute composante connexe de ∂M sur T est triviale.*

Démonstration du Corollaire 4.1.5.— Le théorème de Skora [Sko] (cf. [Ka, Ch. 12], [Ota1, Ch. 8]) montre qu'il existe une sous-surface compacte maximale $S \subset \partial M$ dont le groupe fondamental contient, à conjugaison près dans $\pi_1(M)$, tous les lacets $\gamma \subset \partial M$ qui agissent trivialement sur T. En particulier S contient P.

La maximalité de la surface S et les lemmes 4.1.6 et 4.1.7 ci-dessous montrent que, par une homotopie du bord, S peut être rendue globalement invariante par l'involution de recollement τ.

Le lemme classique suivant (cf. [Mor2, Lemme 11.8]) découle de la définition de la métrique de Teichmüller sur $\mathcal{T}(\partial_0 M)$:

4.1.6. Lemme

Soit (M, P) une variété apprêtée hyperbolique, à bord strictement incompressible. Soient ρ_1 et ρ_2 deux points dans $\mathcal{GF}(M, P)$, tels que $\Delta(\partial\rho_1, \partial\rho_2) \leq \kappa$. Alors, pour tout lacet $\gamma \in \pi_1(M)$:

$(\exp -2\kappa)\ell(\rho_1(\gamma)) \leq \ell(\rho_2(\gamma)) \leq (\exp 2\kappa)\ell(\rho_1(\gamma))$. □

Le lemme 4.1.7 est une conséquence du lemme précédent et du fait que l'application $\tau^\star \circ \sigma$ contracte strictement les distances (lemme 3.2).

4.1.7. Lemme

Il existe une constante $0 < \kappa \leq \Delta(s, \tau^\star \circ \sigma(s))$, telle que pour toute représentation $\rho \in \partial^{-1}\mathcal{O}(s)$ et pour tout lacet $\gamma \subset \partial M$:

$(\exp -2\kappa)\ell(\rho(\tau \circ \gamma)) \leq \ell(\rho(\gamma)) \leq (\exp 2\kappa)\ell(\rho(\tau \circ \gamma))$. □

Par passage à la limite, le lemme 4.1.7 implique pour un lacet $\gamma \subset \partial M$ l'inégalité suivante : $(\exp -2\kappa)\ell_T(\tau \circ \gamma) \leq \ell_T(\gamma) \leq (\exp 2\kappa)\ell_T(\tau \circ \gamma)$.

D'où $\tau(S)$ peut être déformée dans S, et est isotope à S par maximalité de S.

Si la surface S n'est pas tout ∂M, d'après le théorème 4.1.3 de Morgan et Shalen et le lemme 4.1.7, son bord ∂S est une famille finie, non vide, de courbes essentielles sur ∂M, qui est globalement invariante par τ et contenue dans l'intérieur $\mathrm{int}(\partial_0 W)$ du bord des fenêtres W de (M, P).

Le fait que le bord ∂S de la surface S soit invariant (à homotopie près dans ∂M) par une homotopie globale de M permet de construire des anneaux ou des bandes de Möbius essentiels (peut-être singuliers), non parallèles à P et s'appuyant sur les composantes de ∂S. Ces anneaux ou bandes de Möbius s'obtiennent en poussant les courbes de ∂S à travers les fenêtres. Puisque ∂S est globalement invariant par τ, ils se referment en des tores ou des bouteilles de Klein essentiels (peut-être singuliers) dans la variété quotient M/τ, contredisant ainsi son atoroïdalité. □

Ce corollaire montre que l'action du groupe fondamental de chaque composante connexe de $\overline{M - \partial_1 W}$ est triviale sur T. C'est déjà vrai pour les composantes connexes de $\overline{M - W}$ par le théorème 4.1.3 de Morgan et Shalen. C'est vrai pour une fenêtre, car c'est un fibré en intervalles et que l'action du groupe fondamental d'une composante de son bord sur T est triviale et factorise par l'action du groupe fondamental de la fenêtre.

Le fait que l'action du groupe $\pi_1(M)$ est triviale sur T se démontre alors par récurrence. On utilise la structure de graphe de groupes de $\pi_1(M)$, associée au scindage de M le long des anneaux essentiels de $\partial_1 W$, dont tous les stabilisateurs des sommets ont une action triviale sur T. Pour plus de détails nous renvoyons à [Ka, §, 14.17]. □

Nous allons maintenant décrire la dernière étape de la preuve du théorème du point fixe selon Thurston.

4.2. Preuve du théorème de la limite géométrique

Nous commençons par rappeler la définition de la convergence géométrique.

Définition.— *Une suite de variétés hyperboliques pointées* $\{(N_n, x_n)\}_{n\in\mathbb{N}}$ *converge géométriquement vers une variété hyperbolique pointée* (N_∞, x_∞) *si, pour tout* $R > 0$ *et pour tout* $\varepsilon > 0$, *il y a un entier* n_0 *tel que, pour* $n > n_0$, *il existe une application* $(1+\varepsilon)$*-bilipschitz* $f_n : B(x_\infty, R) \to C_n$ *qui vérifie* :

(1) $d(f_n(x_\infty), x_n) < \varepsilon$,
(2) $B(x_n, R-\varepsilon) \subset f_n(B(x_\infty, R)) \subset B(x_n, R(1+\varepsilon)+\varepsilon)$.

En général, l'holonomie $\rho_\infty \in \mathcal{X}(\pi_1(N_\infty))$ de la variété hyperbolique limite est différente de la limite simple des holonomies des variétés hyperboliques N_n. Cela reste cependant vrai sur les compacts de N_∞.

4.2.1. Lemme

Soit $\{(N_n, x_n)\}_{n\in\mathbb{N}}$ *une suite de variétés hyperboliques pointées qui converge géométriquement vers une variété hyperbolique pointée* (N_∞, x_∞). *Soit* ρ_n *(resp.* ρ_∞*) l'holonomie de la variété hyperbolique* N_n *(resp.* N_∞*). Soit* $A \subset N_\infty$ *un compact contenant le point base* x_∞. *Alors pour tout* $n \in \mathbb{N}$, *il existe un plongement* $(1+\varepsilon_n)$*-quasi-isométrique* $f_n : A \to N_n$, *avec* $\lim_{n\to\infty} \varepsilon_n = 0$ *et tel que* : $\forall \gamma \in \pi_1(A, x_\infty)$,

$$\lim_{n\to\infty} \rho_n(f_{n*}(\gamma)) = \rho_\infty(\gamma).$$

□

Le théorème de compacité suivant, dû à Gromov dans le cadre plus général des variétés riemanniennes à courbure sectionnelle pincée, est le point de départ de la preuve du théorème 4.2 de la limite géométrique.

4.2.2. Théorème de compacité

Soit $\{(N_n, x_n)\}_{n\in\mathbb{N}}$ *une suite de variétés hyperboliques pointées, dont les points bases ont un rayon d'injectivité uniformément minoré (par exemple par la constante de Margulis). Alors, il existe une suite extraite* $\{(N_{n_k}, x_{n_k})\}_{n_k\in\mathbb{N}}$ *qui converge géométriquement vers une variété hyperbolique* (N_∞, x_∞). □

On peut trouver une preuve de ce théorème, pour le cas particulier énoncé, dans [CEG] et [Th0].

Soit ρ_∞ un point d'accumulation de $\overline{\partial^{-1}(\mathcal{O}(s))}$. Une conséquence du lemme de Margulis et du théorème de compacité de Gromov est l'existence d'une suite de représentations $\{\rho_k\}_{k\in\mathbb{N}} \subset \partial^{-1}(\mathcal{O}(s))$, qui converge vers ρ_∞ dans $\mathcal{X}(M, P)$ et telle que la suite de variétés hyperboliques pointée $\{(\mathbb{H}^3/\rho_k(\pi_1 M), x_k)\}_{k\in\mathbb{N}}$ converge géométriquement vers une variété hyperbolique pointée (Q_∞, x_∞) (pour une suite de points bases $\{x_k\}_{k\in\mathbb{N}}$ convenablement choisis).

Dans toute la suite, nous noterons N_k et N_∞ les variétés hyperboliques $(\mathbb{H}^3/\rho_k(\pi_1 M))$ et $(\mathbb{H}^3/\rho_\infty(\pi_1 M))$. Nous dirons que la variété hyperbolique N_∞ est la *limite algébrique* de la suite de variétés hyperboliques $\{N_k\}_{k\in\mathbb{N}}$.

Le lemme 4.2.1 montre qu'après une conjugaison convenable dans $PSL_2(\mathbb{C})$, on peut supposer que $\pi_1(N_\infty) = \rho_\infty(\pi_1(M))$ est un sous-groupe du groupe $G = \pi_1(Q_\infty) \subset PSL_2(\mathbb{C})$. En fait, on peut identifier le groupe G à *l'enveloppe* de la suite $\{\rho_k(\pi_1(M))\}_{k\in\mathbb{N}}$ (cf. [JM], [MT, Ch. 7]) :

$G = \{g = \lim_{k\to\infty} g_k \in PSL_2(\mathbb{C}), g_k \in \rho_k(\pi_1(M))\}$.

La limite algébrique N_∞ est donc un revêtement localement isométrique de la limite géométrique Q_∞. Pour démontrer le théorème de la limite géométrique il suffit de montrer que ce revêtement est de degré fini, grâce au lemme suivant :

4.2.3. Lemme

Avec les notations ci-dessus, $\rho_\infty(\pi_1(M)) = G$ *si et seulement si* $\rho_\infty(\pi_1(M))$ *est d'indice fini dans* G.

Démonstration du lemme 4.2.3.— Soit $g = \lim_{k\to\infty} \rho_k(\gamma_k) \in G$, avec $\gamma_k \in \pi_1(M)$. S'il existe un entier $n \in \mathbb{N}$ tel que $g^n = \rho_\infty(\gamma)$, avec $\gamma \in \pi_1(M)$, alors $\lim_{k\to\infty} \rho_k(\gamma^{-1}\gamma_k^n) = 1$. Comme G est un sous-groupe discret de $PSL_2(\mathbb{C})$, $\gamma^{-1}\gamma_k^n = 1$ pour presque tout $k \in \mathbb{N}$. D'où γ_k est la racine k-ième de γ pour presque tout $k \in \mathbb{N}$ et $g \in \rho_\infty(\pi_1(M))$. □

Jusqu'à maintenant nous n'avons considéré que des variétés hyperboliques géométriquement finies. La difficulté de la preuve du théorème de la limite géométrique réside dans le fait qu'a priori la limite algébrique N_∞ n'est pas géométriquement finie, pas plus que la limite géométrique Q_∞, dont on ne contrôle même pas le groupe fondamental.

Comme le groupe $\pi_1(N_\infty)$ est de type fini, N_∞ admet un cœur compact relatif $M_\infty \subset N_\infty$. C'est une sous-variété compacte dont l'inclusion est une équivalence d'homotopie, et telle que la trace sur ∂M_∞ de la partie cuspidale N_∞^{cusp} est une réunion disjointe d'anneaux et de tores que l'on note P_∞.

Par abus de langage, on appellera *bout non cuspidal* de N_∞ l'adhérence d'une composante connexe de $N_\infty - \{M_\infty \cup N_\infty^{cusp}\}$. Les bouts non cuspidaux de N_∞ sont alors en bijection avec les composantes connexes de $\partial_0 M_\infty = \overline{\partial M_\infty - P_\infty}$.(cf. [Bon]).

Un bout non cuspidal $E \subset N_\infty$ est dit *géométriquement fini* si son intersection avec le cœur convexe $C(N_\infty)$ est compacte. Sinon il est dit *géométriquement infini.*

Un bout non cuspidal E, associé à une composante $F \subset \partial_0 M_\infty$, est dit *simplement dégénéré* si pour tout compact $K \subset E$ il existe une courbe fermée simple $\beta \subset F$ qui soit homotope dans N_∞ à une géodésique fermée $\gamma \subset E - K$. C'est une façon très restrictive d'être géométriquement infini. En particulier, un tel bout est difféomorphe au produit $F \times [0, +\infty[$.

Puisque, par hypothèse, la surface $\partial_0 M$ est incompressible dans M, la surface $\partial_0 M_\infty$ est incompressible dans M_∞. Le groupe kleinien $\pi_1(N_\infty)$ vérifie donc la condition de Bonahon [Bon]. Le théorème de Bonahon montre que N_∞ est difféomorphe à l'intérieur de la variété compacte M_∞. En particulier chaque bout non cuspidal de N_∞ est difféomorphe au produit $F \times [0, +\infty[$, pour une composante connexe $F \subset \partial_0 M_\infty$. Plus précisément, F. Bonahon a démontré le théorème suivant :

4.2.4. Théorème des bouts ([Bon])

Soit N une variété hyperbolique, dont le groupe fondamental est de type fini, non cyclique et possède la propriété suivante : pour toute décomposition en produit libre de $\pi_1(N)$, il existe un élément parabolique qui n'est conjugué dans aucun des facteurs. Alors, tout bout géométriquement infini de N est simplement dégénéré. □

Pour plus de détails sur les bouts des variétés hyperboliques et la preuve du théorème des bouts nous renvoyons à [Bon].

Pour achever la preuve du théorème de l'orbite bornée on doit montrer que la limite algébrique N_∞ est géométriquement finie, c'est-à-dire que ses bouts non cuspidaux sont géométriquement finis.

Le théorème suivant de Thurston permet de contrôler le revêtement $q : N_\infty \to Q_\infty$ sur les bouts géométriquement infinis de N_∞.

4.2.5. Théorème du revêtement ([Th0, Ch. 9])

Soit $p : N \to Q$ un revêtement localement isométrique d'une variété hyperbolique Q de volume infini. On suppose que $\pi_1(N)$ est de type fini, sans parabolique accidentel et qu'il vérifie la condition de Bonahon. Pour tout bout géométriquement infini $E \subset N$, il existe un compact $K \subset E$ tel que la restriction de l'application de revêtement p à $E - K$ est propre et de degré fini.

Remarque : Il existe un énoncé plus général, dû à R. Canary [Ca].

Démonstration du théorème du revêtement (cf. [Th0, Thm. 9.2.2], [Oh, Lemme 2.2]) : Soit $M \subset N - N^{cusp}$ le cœur compact de N. Soit $E \subset N - \{M \cup N^{cusp}\}$ un bout non cuspidal et géométriquement infini, correspondant à une composante $F \subset \partial_0 M$. On note encore q la restriction de l'application de revêtement au bout E. Comme ce bout est difféomorphe au produit $F \times [0, +\infty[$, cette restriction est de degré fini si et seulement si $q^{-1}(q(F))$ est compact. On raisonne alors par contradiction en supposant $q^{-1}(q(F))$ non compact. Deux cas peuvent se produire :

a) Il existe une composante $\Sigma \subset q^{-1}(q(F))$ qui n'est pas bornée dans E.

b) L'ensemble $q^{-1}(q(F))$ admet une infinité de composantes connexes compactes, qui sont des surfaces incompressibles immergées dans E.

Dans le premier cas a), puisque $\pi_1(F)$ n'admet pas de parabolique accidentel, le théorème des bouts montre qu'il existe une suite de courbes simples fermées $\beta_n \subset F$ qui sont homotopes à des géodésiques $\gamma_n \subset E$, telles que toute suite extraite non stationnaire de $\{\gamma_n\}_{n\in\mathbb{N}}$ s'échappe de tout compact de E et que les courbes γ_n restent à distance uniformément bornée de Σ. D'où la suite de géodésiques $q(\gamma_n)$ reste à distance uniformément bornée de la surface compacte immergée $q(F) \supset q(\Sigma)$. Comme le normalisateur de $\pi_1(F)$ dans $\pi_1(Q)$ est réduit à lui-même, il existe une suite extraite de la suite $q(\gamma_n)_{n\in\mathbb{N}}$ qui converge vers une géodésique fermée $\bar{\gamma}$, homotope dans Q à l'image $q(\beta)$ d'une courbe simple fermée sur $\beta \subset F$. Il existe alors une suite extraite de $\{\gamma_n\}_{n\in\mathbb{N}}$ convergeant dans N vers une géodesique fermée homotope à β, ce qui contredit le choix de la suite $\{\gamma_n\}_{n\in\mathbb{N}}$.

Dans le second cas b), $q^{-1}(q(F))$ admet des composantes S_1 et S_2 qui sont des surfaces immergées, disjointes et incompressibles dans E. Comme E est difféomorphe au produit $F \times [0, +\infty[$, une fois les surfaces F, S_1 et S_2 orientées, il existe des entiers $k, \ell, m \in \mathbb{Z}$ tels que les cycles $k[F]$, $\ell[S_1]$ et $m[S_2]$ soient homologues dans $\mathcal{Z}_2(E, E \cap N^{cusp})$. Il existe donc une 3-chaîne $C \subset \mathcal{C}_3(E, E \cap N^{cusp})$ telle que $\partial C = \ell[S_1] - m[S_2] \in \mathcal{Z}_2(E, E \cap N^{cusp})$. D'où $\partial q_*(C) = 0$ dans $\mathcal{Z}_2(Q - Q^{cusp}, \partial(Q - Q^{cusp}))$, et $q_*(C)$ est un cycle non trivial dans $\mathcal{Z}_3(Q - Q^{cusp}, \partial(Q - Q^{cusp}))$. La variété $Q - Q^{cusp}$ serait alors compacte et Q aurait donc un volume fini contrairement à l'hypothèse. □

Nous allons maintenant montrer comment déduire le théorème de convergence géométrique du théorème du revêtement.

On peut remarquer d'abord que la variété hyperbolique Q_∞ a un volume infini, comme limite géométrique de variétés hyperboliques de volume infini.

L'hypothèse d'atoroïdalité de la variété quotient M/τ va permettre d'éviter les paraboliques accidentels dans $\rho_\infty(\pi_1(M)) = \pi_1(N_\infty)$.

4.2.6. Lemme

Avec les hypothèses du théorème de la limite géométrique, le groupe kleinien $\rho_\infty(\pi_1 M)$ n'admet pas de parabolique accidentel.

Démonstration du lemme 4.2.6 : L'ensemble des paraboliques accidentels de $\rho_\infty(\pi_1 M)$ correspond à une famille finie $\mathcal{P}$ de courbes simples, fermées et disjointes sur $\partial_0 M$, qui ne sont pas parallèles entre elles, ni parallèles à ∂P. Ceci découle de l'étude des cusps accidentels dans le revêtement hyperbolique N_∞^F de N_∞, associé à une composante connexe $F \subset \partial_0 M$ (cf. [Ka, § 6.21], [Mor1, Ch. 12]).

Ces courbes $\gamma \in \mathcal{P}$ sont caractérisées par le fait que : $\lim_{k\to\infty} \ell(\rho_k(\gamma)) = 0$.

En particulier, toute courbe simple fermée sur $\partial_0 M$, qui est homotope dans M à une courbe de $\mathcal{P}$, est parallèle sur $\partial_0 M$ à une courbe de $\mathcal{P}$.

Si l'on suppose que l'ensemble $\mathcal{P}$ est non vide, le lemme 4.1.7 montre que l'involution τ laisse globalement invariant cet ensemble : $\tau(\mathcal{P}) = \mathcal{P}$.

Le lemme 4.2.7, ci-dessous, montre que l'on peut pousser les courbes de $\mathcal{P}$ dans la sous-variété caractéristique de M. En particulier, on peut supposer que $\mathcal{P} \subset \{\partial_0 W \cup A\}$, où $\partial_0 W \subset \partial_0 M$ est le bord des fenêtres W et $A \subset \partial_0 M$ est une réunion d'anneaux, appartenant au bord des composantes connexes de $\overline{M - W}$ qui sont des tores solides.

On utilise alors la structure de I-fibré des fenêtres W et les anneaux essentiels de $A \cap \partial_1 W$ pour construire un tore (ou une bouteille de Klein) essentiel (peut-être singulier) dans la variété quotient M/τ, contredisant ainsi son atoroïdalité. □

4.2.7. Lemme

Soit (M, P) une variété apprêtée hyperbolique, à bord strictement incompressible et qui n'est pas un fibré en intervalles. Soit un point $s \in \mathcal{T}(\partial_0 M)$. Il existe une constante $\eta > 0$, ne dépendant que du point s et du type topologique (M, P), telle que tout élément non parabolique $\gamma \in \pi_1(\partial_0 M)$, ayant une longueur de translation $\ell(\rho(\gamma)) < \eta$ pour une représentation $\rho \in \partial^{-1}(\mathcal{O}(s)$, est conjugué dans le groupe fondamental d'une composante connexe de la sous-variété caractéristique de M. □

La preuve de ce lemme est assez technique et délicate. Elle utilise les tubes de Margulis, associés aux lacets de longueur suffisamment petite sur $\partial_0 M$, dans les revêtements quasi-fuchsiens correspondant aux composantes de bord qui les portent. Une analyse précise de ces tubes permet de construire des anneaux essentiels (peut-être singuliers) dans M, s'appuyant sur ces courbes dans $\partial_0 M$. On peut alors pousser ces anneaux dans la sous-variété caractéristique de M. Pour une exposition détaillée de cette construction nous renvoyons à [Ka, § 17.2.3], [Mor1, § 12].

Nous pouvons maintenant expliquer la preuve de la première assertion du Théorème de la limite géométrique (cf. [Th0, Thm. 9.6.1.a], [Oh, Thm. 2.1]).

Démonstration de l'assertion 1) *du théorème* 4.2 : Soit $q : N_\infty \to Q_\infty$ le revêtement de la limite géométrique par la limite algébrique. On veut montrer que son degré est fini.

Une surface plissée dans une variété hyperbolique N est une application $f : S \to N$ d'une surface hyperbolique complète S dans N, telle que :

(i) tout point $x \in S$ appartient à l'intérieur d'un arc géodésique dont l'image par f est un arc géodésique de N,

(ii) l'application f est injective au niveau des groupes fondamentaux.

Dans ce qui suit, nous identifierons une surface plissée avec son image $f(S)$ dans N. En particulier, le bord du cœur convexe de N est constitué de surface plissée, lorsque N vérifie la condition de Bonahon.

Soit $F \subset \partial_0 M$ une composante connexe. Comme le bout de la variété hyperbolique N_k correspondant à $\rho_k(\pi_1(F))$ est géométriquement fini, il existe une surface plissée $S_k \subset \partial C(N_k)$ dans le bord du cœur convexe de N_k, telle qu'une composante connexe de $N_k - S_k$ ne rencontre pas $C(N_k)$. Une telle surface est unique et induit une structure hyperbolique marquée sur F. On peut donc lui associer un point $[S_k]$ dans l'espace de Teichmüller $\mathcal{T}(F)$. Quitte à extraire une sous-suite, on peut supposer que la suite des points $\{[S_k]\}_{k\in\mathbb{N}}$ converge dans la compactification de Thurston $\overline{\mathcal{T}(F)}$ de l'espace de Teichmüller de F. On rappelle que $\overline{\mathcal{T}(F)} = \mathcal{T}(F) \cup \mathcal{PL}(F)$, où $\mathcal{PL}(F)$ est l'espace projectif des laminations mesurées sur F (cf. [FLP], [Th7]). On distingue alors deux cas :

a) la suite $\{[S_k]\}_{k\in\mathbb{N}}$ converge dans $\mathcal{T}(F)$. On peut réaliser la limite par une surface plissée $S_\infty \subset \partial C(N_\infty)$ dans le bord du cœur convexe de N_∞, qui correspond à $\rho_\infty(\pi_1(F))$ (cf. [CEG], [EM]). Comme $\rho_\infty(\pi_1(F))$ n'admet pas de parabolique accidentel, le théorème de Thurston de compacité des surfaces plissées (cf. [Th0, §, 8.5], [CEG, Thm 5.2]) permet de montrer que l'image $q(S_\infty) \subset \partial C(Q_\infty)$ est la limite géométrique des surfaces plissées $S_k \subset \partial C(N_k)$. On en déduit que le bout non cuspidal E_∞ de N_∞, associé à S_∞, est géométriquement fini. De plus la restriction de l'application de revêtement q à ce bout est un plongement. En effet, si $\tilde{S}_\infty \subset \mathbb{H}^3$ est un relevé de S_∞, la convergence géométrique des variétés N_k (respectivement des surfaces plissées S_k) vers Q_∞ (respectivement $q(S_\infty)$) implique que les translatés de $\tilde{S}_\infty$ par les éléments de $\pi_1(Q_\infty)$, qui ne normalisent pas $q_\star(\pi_1(S_\infty))$, sont disjoints. Or par hypothèse, le normalisateur de $q_\star(\pi_1(S_\infty))$ dans $\pi_1(Q_\infty)$ est lui-même ;

b) la suite $\{[S_k]\}_{k\in\mathbb{N}}$ converge dans $\overline{\mathcal{T}(F)}$ vers une lamination mesurée $\mathcal{L}_\infty$. Comme $\rho_\infty(\pi_1(F))$ n'admet pas de parabolique accidentel, le théorème de Thurston de compacité des surfaces plissées (ayant un type d'homotopie fixé) (cf. [Th0, § 8.5], [CEG, Thm 5.2.18]), montre que $\mathcal{L}_\infty$ est la *lamination terminale* d'un bout géométriquement infini du revêtement N_∞^F de N_∞, associé au sous-groupe $\rho_\infty(\pi_1(F))$. D'après le théorème du revêtement, $\rho_\infty(\pi_1(F))$ est d'indice fini dans le groupe fondamental $\pi_1(E_\infty)$ du bout correspondant $E_\infty \subset N_\infty$.

La discussion des deux cas précédents montre que le bord du cœur compact M_∞ de N_∞ a les propriétés suivantes :

(i) le groupe fondamental de chaque composante connexe de $\partial_0 M_\infty$ contient, comme sous-groupe d'indice fini, l'image par ρ_∞ du groupe fondamental d'une composante connexe de $\partial_0 M$;

(ii) deux composantes connexes de $\partial_0 M_\infty$ ne sont jamais homotopes dans N_∞, car le groupe fondamental $\pi_1(N_\infty) = \pi_1(M)$ ne contient pas un sous-groupe de surface d'indice fini.

Comme M et M_∞ sont des variétés irréductibles, ayant le même type d'homotopie, leurs bords ont la même caractéristique d'Euler. Les propriétés (i) et (ii) montrent alors que, pour chaque composante connexe $F_\infty \subset \partial_0 M_\infty$, il existe une unique composante connexe $F \subset \partial_0 M$ telle que $\rho_\infty(\pi_1(F)) = \pi_1(F_\infty)$. En particulier l'isomorphisme $\rho_\infty : \pi_1(M) \to \pi_1(M_\infty)$ préserve la structure périphérique et peut donc être réalisé par un difféomorphisme d'après le théorème de F. Waldhausen [Wa].

Il en découle que la restriction de l'application de revêtement q sur les bouts non cuspidaux de N_∞ est de degré fini, car elle est de degré 1 (d'après ce qui précède) sur les bouts géométriquement finis et de degré fini sur les bouts géométriquement infinis (d'après le théorème du revêtement).

Puisque la partie non cuspidale de N_∞ est réunion de son cœur compact M_∞ et des bouts non cuspidaux, la restriction de q à $N_\infty - N_\infty^{cusp}$ est de degré fini. Comme q envoie la partie cuspidale de N_∞ sur la partie cuspidale de Q_∞, $\pi_1(N_\infty)$ est d'indice fini dans $\pi_1(Q_\infty)$ et le lemme 4.2.3 permet de conclure. □

Nous achevons maintenant la preuve du théorème de la limite géométrique en démontrant l'assertion 2). Ceci achèvera donc la preuve du théorème du point fixe 3.1, et par là même la preuve du théorème de recollement 2.1.

Démonstration de l'assertion 2) *du théorème* 4.2 (cf. [Ka, §, 17.2.4]) : Le but est de montrer, en utilisant l'assertion 1), que tous les bouts non cuspidaux de la limite algébrique N_∞ sont géométriquement finis.

Soient $E^+ \subset N_\infty$ un bout non cuspidal, correspondant à une composante $F^+ \subset \partial_0 M$ et $E^- \subset N_\infty$ le bout correspondant à la composante $F^- = \tau(F^+) \subset \partial_0 M$. On considère les revêtements $p_+ : \hat{N}^+ \to N_\infty$ et $p_- : \hat{N}^- \to N_\infty$ respectivement associés aux sous-groupes $\pi_1(F^+)$ et $\pi_1(F^-)$ de $\pi_1(N_\infty)$.

Ces revêtements ont respectivement une paire de bouts $(\hat{E}_0^+, \hat{E}_1^+)$ et $(\hat{E}_0^-, \hat{E}_1^-)$. Du fait que l'isomorphisme $\rho_\infty : \pi_1(M) \to \pi_1(N_\infty)$ préserve la structure périphérique, la restriction des applications de revêtements p^+ et p^- à un des bouts de chaque paire, par exemple $\hat{E}_0^+$ et $\hat{E}_0^-$, est une isométrie.

De plus, si $\tau = (f, f^{-1}) : F^+ \sqcup F^- \to F^- \sqcup F^+$, la convergence géométrique implique que l'application d'identification $f : F^+ \to F^-$ induit un homéomorphisme quasi-conforme au bord $\hat{f}_\infty : \hat{N}^+ \to \hat{N}^-$ qui transpose les paires de bouts : $\hat{f}_\infty(\hat{E}_0^+) = \hat{E}_1^-$ et $\hat{f}_\infty(\hat{E}_1^+) = \hat{E}_0^-$.

Si le bout non cuspidal E^+ de N_∞ est géométriquement fini, le revêtement $\hat{N}^+$ est quasifuchsien, d'où le revêtement $\hat{N}^-$ est quasifuchsien, et le bout E^- est lui aussi géométriquement fini. En particulier, si le bout E^+ est géométriquement infini, les bouts $\hat{E}_0^+ \subset \hat{N}^+$ et $\hat{E}_0^- \subset \hat{N}^-$ sont géométriquement infinis. Puisque $\hat{f}_\infty(\hat{E}_1^+) = \hat{E}_0^-$, le bout

$\hat{E}_1^+ \subset \hat{N}^+$ est lui aussi géométriquement infini. Le théorème du revêtement montrerait alors que p^+ est un revêtement fini, contredisant l'hypothèse que M et donc N_∞ n'est pas un fibré en intervalles. □

5. VARIÉTÉS MIROITÉES

Nous allons maintenant décrire la partie topologique de la récurrence, qui permet de ramener la preuve du théorème d'hyperbolisation à celle du théorème de recollement final. Il s'agit de la partie la plus classique de la preuve du théorème d'hyperbolisation. C'est dans cette étape qu'on utilise l'existence d'une *hiérarchie finie* pour une variété hakenienne de dimension 3. Cela permet de lui associer une complexité et donc de raisonner par récurrence sur cette complexité. Nous suivons ici l'approche adoptée dans [Ota2, § 7], auquel nous référons pour les détails.

Dans cette étape, il est nécessaire de considérer des hiérarchies adaptées à la notion de variétés apprêtées. Ce type de hiérarchie avait déjà été considéré par K. Johannson [Jo] dans la preuve de son théorème sur les équivalences d'homotopie et la finitude du groupe des difféotopies d'une variété hakenienne atoroïdale.

Définition.— *Une surface essentielle* $(F, \partial F) \hookrightarrow (M, \partial M)$ *est dite* strictement essentielle *dans la variété apprêtée* (M, P) *si elle vérifie les conditions suivantes* :

1) ∂F *est transverse à* ∂P *et minimise le cardinal de* $\partial F \cap \partial P$ *dans sa classe d'isotopie* ;

2) *toute composante connexe de* ∂F *parallèle à* ∂P *appartient à* P ;

3) *il n'existe pas d'anneau incompressible plongé* $(S^1 \times [0,1], S^1 \times \{0\}, S^1 \times \{1\}) \hookrightarrow (M, \partial M, F)$, *tel que la composante de bord* $S^1 \times \{1\} \subset F$) *ne soit pas parallèle à* ∂F.

La proposition ci-dessous décrit une notion de hiérarchie qui est une extension naturelle aux variétés apprêtées de la construction classique d'une hiérarchie pour une variété hakenienne (cf. [Jo], [Mor1, § 4], [Ota2, § 7]).

5.1. Proposition

Si M *est une variété hakenienne, toute variété apprêtée atoroïdale* (M,P) *admet une hiérarchie du type suivant : il existe une suite finie* $(M_0, P_0), \ldots, (M_n, P_n)$ *de variétés apprêtées atoroïdales telles que* :

(i) $(M_0, P_0) = (M, P)$;

(ii) *pour* $k \leq n-1$, *il existe une surface connexe strictement essentielle* $F_k \subset M_k$ *(qui n'est pas un disque et dont le bord est non vide si* $\partial_0 M_k \neq \emptyset$*) telle que* M_{k+1} *est obtenue en coupant* M_k *le long de* F_k. *De plus,* P_{k+1} *est obtenu en coupant* P_k *le long de* ∂F_k *et en oubliant les composantes qui sont des disques* ;

(iii) M_n *est une réunion de corps à anses (ou bretzels).*

On appelle n la longueur de la hiérarchie. L'existence d'une telle hiérarchie et le fait que sa longueur soit bornée par une constante ne dépendant que du type topologique de M résultent des deux lemmes classiques suivants :

5.2. Lemme de finitude de Haken-Kneser

Dans une variété compacte orientable irréductible de dimension 3, il existe au plus un nombre fini $h(M)$ de surfaces essentielles, disjointes et deux à deux non parallèles. □

5.3. Lemme

Une variété hakenienne est un corps à anses si et seulement si toute surface essentielle dans M est un disque. □

On définit alors la *complexité de la variété apprêtée (M,P)* comme la paire $c(M,P) = (\ell(M,P), g(\partial M)) \subset \mathbb{N} \times \mathbb{N}$, où $\ell(M,P)$ est la plus grande longueur possible pour une hiérarchie de la variété apprêtée (M,P), donnée par la proposition 5.1, et $g(\partial M)$ est le genre du bord de M. On munit $\mathbb{N} \times \mathbb{N}$ de l'ordre lexicographique.

Une variété de complexité minimale est alors une réunion disjointe de boules $(B^3, \emptyset)$. De plus la complexité d'une variété apprêtée **décroît strictement** lorsqu'on la découpe le long d'une surface strictement essentielle.

La jolie idée de Thurston est d'associer à une variété atoroïdale apprêtée (M,P) une variété atoroïdale *miroitée*, à bord vide ou constitué de tores.

Définitions.— *On appelle variété miroitée un couple $(\hat{M}, H)$ formé d'une variété compacte orientable irréductible $\hat{M}$, à bord vide ou constitué de tores incompressibles, et d'un sous-groupe abélien fini $H = (\mathbb{Z}_2)^k \subset Diff(\hat{M})$ engendré par des symétries par rapport à des surfaces essentielles dans $\hat{M}$.*

On appelle type topologique de la variété miroitée $(\hat{M}, H)$ la variété apprêtée (M,P) où M est l'espace sous-jacent du quotient $\hat{M}/H$ et P est l'espace sous-jacent du quotient $\partial\hat{M}/H$ dans lequel on oublie les composantes qui sont des disques.

Remarque sur les kaléidoscopes : La structure de variété miroitée $(\hat{M}, H)$ est déterminée par la stucture *orbifold* du quotient $\mathcal{K} = \hat{M}/H$. C'est la structure d'espace métrique sur $\mathcal{K} = \hat{M}/H$ dans laquelle tout point x a un voisinage localement modelé sur le quotient de $\mathbb{R}^3$ par un sous-groupe du groupe à huit éléments engendré par les symétries par rapport aux trois plans de coordonnées. On appellera ce sous-groupe le groupe d'isotropie du point $x \in \mathcal{K}$ considéré. Le bord orbifold $\partial\mathcal{K}$ de $\mathcal{K}$ correspond à l'orbifold quotient $\partial\hat{M}/H$. On appelle lieu singulier de l'orbifold $\mathcal{K}$ l'ensemble $\Sigma \subset \partial_0 M$ des points ayant un groupe d'isotropie non trivial. Dans notre cas c'est une sous-orbifold de dimension 2, dont l'espace sous-jacent est le complémentaire dans ∂M de l'espace sous-jacent du bord $\partial\mathcal{K}$.

L'ensemble des points, ayant un groupe d'isotropie d'ordre ≥ 4, est une sous-orbifold $\mathcal{G} \subset \Sigma$ de dimension 1. Son espace sous-jacent est un graphe trivalent dans $\partial_0 M$, dont

les sommets correspondent aux points ayant un groupe d'isotropie maximal dans $\mathcal{K}$. On appelle l'adhérence des composantes connexes de $\Sigma - \mathcal{G}$ les *facettes* de $\mathcal{K}$.

On peut donc voir l'orbifold $\mathcal{K}$ comme un *kaléidoscope* à angles droits. Il est obtenu à partir de la variété apprêtée (M, P) en ciselant des facettes sur $\partial_0 M$ le long des composantes connexes de $\partial_0 M - \mathcal{G}$, dont le bord n'est pas un rectangle dans ce graphe.

Cette notion de kaléidoscope quotient est cruciale pour établir les propriétés utiles des variétés miroitées. Le fait que pour la preuve du théorème d'hyperbolisation on ait besoin de considérer seulement ce type de kaléidoscopes à angles droits est une observation de F. Bonahon.

5.4. Proposition

Soit (M, P) une variété apprêtée atoroïdale, l'ensemble $\mathcal{AM}(M, P)$ des variétés miroitées atoroïdales de type topologique (M, P) est non vide.

Démonstration de la proposition 5.4 : D'après la remarque précédente, la preuve de la proposition 5.4 consiste essentiellement à trouver un graphe trivalent $\mathcal{G}$ dans $\partial_0 M$ ayant les propriétés suivantes :

(i) l'adhérence de toute composante connexe du complémentaire $\partial_0 M - \mathcal{G}$ est soit un n-gone ayant $n \geq 5$ sommets, soit un anneau avec au moins un sommet de $\mathcal{G}$ dans une composante de bord, l'autre composante de bord étant incluse dans ∂P ;

(ii) toute courbe simple fermée $\gamma \subset \partial_0 M$, qui rencontre transversalement le graphe $\mathcal{G}$ en au plus trois points et qui borde un disque dans M, borde un disque D dans $\partial_0 M$ dont l'intersection $D \cap \mathcal{G}$ est le cône sur $\partial D \cap \mathcal{G}$.

(iii) toute courbe simple fermée $\gamma \subset \partial_0 M$, qui rencontre transversalement le graphe $\mathcal{G}$ en quatre points et qui borde un disque dans M, borde un disque D dans $\partial_0 M$ dont l'intersection $D \cap \mathcal{G}$ consiste en deux arcs disjoints d'extrémités $\partial D \cap \mathcal{G}$, avec au plus deux arêtes parallèles de $\mathcal{G}$ les reliant dans D.

Le kaléidoscope $\mathcal{K}$ associé, dont les facettes sont obtenues en ciselant toutes les composantes connexes de $\partial_0 M - \mathcal{G}$ est irréductible et atoroïdal au sens des orbifolds (cf. [BS]). Dans notre cas c'est équivalent, via le théorème du lacet et de la sphère équivariants ([DD], [MY]), à l'irréductibilité et l'atoroïdalité de la variété $\hat{M}$, obtenue par miroitage du kaléidoscope $\mathcal{K}$ le long de ses facettes.

L'irréductibilité découle de l'irréductibilité de la variété M, de l'incompressibilité de $\partial_0 M$ et des propriétés (i) et (ii) du graphe $\mathcal{G}$. L'atoroïdalité découle de l'atoroïdalité de la variété apprêtée (M, P) et des propriétés (i) et (iii) du graphe $\mathcal{G}$.

Pour obtenir le graphe $\mathcal{G}$ ayant les propriétés voulues, on considère un voisinage ouvert collier de P dans ∂M, et une triangulation de son complémentaire dans ∂M. On modifie alors cette triangulation de la façon suivante : on ajoute 3 sommets correspondant au milieu des côtés du triangle, puis les 3 arêtes (parallèles aux côtés du triangle) qui les joignent deux à deux. On réitère cette construction dans le triangle inscrit et on joint

chacun des sommets du nouveau triangle obtenu, au sommet correspondant du triangle de départ. Le 1-squelette de la cellulation duale de cette triangulation modifiée a les propriétés requises.

Pour les détails de cette construction due à E. Giroux, nous renvoyons à [Ota2, § 7]. □

Définition *On note $\mathcal{HM}(M,P)$ l'ensemble des variétés miroitées $(\hat{M},H)$ de type topologique (M,P) et admettant une structure hyperbolique complète de volume fini, invariante par H. Clairement $\mathcal{HM}(M,P) \subset \mathcal{AM}(M,P)$.*

Lorsque la variété apprêtée M n'est pas virtuellement fibrée sur le cercle, le théorème suivant se démontre, grâce au théorème de recollement 2.1, par récurrence sur la complexité $c(M,P)$ de la variété apprêtée (M,P).

5.5. Théorème d'hyperbolisation des variétés miroitées

Soit (M,P) une variété apprêtée atoroïdale. Si M est une variété hakenienne, $\mathcal{HM}(M,P) = \mathcal{AM}(M,P)$.

Remarque : L'existence d'une structure hyperbolique complète à volume fini, invariante par le groupe des miroirs H, sur la variété miroitée $\hat{M}$ est équivalente à l'existence d'une structure hyperbolique à volume fini sur le kaléidoscope quotient $\mathcal{K} = \hat{M}/H$, ayant les propriétés suivantes :

(i) toutes les facettes de $\mathcal{K}$ sont totalement géodésiques,

(ii) l'angle dièdre de deux facettes le long d'une arête commune de $\mathcal{G}$ est $\pi/2$.

Grâce à la proposition 5.4, le lemme suivant permet de déduire le théorème d'hyperbolisation pour les variétés apprêtées du théorème d'hyperbolisation pour les variétés miroitées.

5.6. Lemme

Soit (M,P) une variété apprêtée. Si $\mathcal{HM}(M,P)$ est non vide, (M,P) est hyperbolique.

Démonstration du lemme 5.6 : S'il existe une variété miroitée hyperbolique $(\hat{M},H)$ de type topologique (M,P), le kaleidoscope $\mathcal{K} = \hat{M}/H$ admet une structure hyperbolique à volume fini et totalement géodésique le long de ses facettes. L'ensemble des points intérieurs et non singuliers de $\mathcal{K}$ s'identifie à l'intérieur d'un convexe de volume fini dans $\mathcal{K}$. Cela permet de munir M d'une structure hyperbolique géométriquement finie. □

Nous décrivons maintenant la preuve du théorème d'hyperbolisation pour les variétés miroitées, dont le type topologique n'est pas virtuellement fibré sur le cercle.

Démonstration du théorème 5.5 *dans le cas non fibré* : L'étape initiale de la récurrence est donnée par le cas particulier suivant du théorème d'Andreev, qui montre que les kaléidoscopes atoroïdaux de type topologique $(B^3,\emptyset)$ sont hyperboliques. Pour une preuve de ce théorème nous renvoyons aux articles originaux de E.M. Andreev [And1, And2], aux notes de Thurston [Th0, § 13] et aux travaux de I. Rivin [Riv] (cf. [RH]).

5.7. Théorème d'Andreev ([And1, And2])

Les trois conditions suivantes caractérisent combinatoirement un polyèdre convexe de volume fini et d'angles dièdres $\pi/2$ dans $\mathbb{H}^3$, qui n'est ni un simplexe, ni un prisme triangulaire :

(1) *trois faces deux à deux adjacentes sont concourantes en un sommet* ;
(2) *deux faces non adjacentes, mais adjacentes à une même face, sont concourantes avec cette face en un sommet, ou sont disjointes* ;
(3) *quatre faces cycliquement adjacentes sont concourantes en un sommet ou trois des faces sont deux à deux adjacentes.* □

Définition.— *Soit $(\hat{M}, H)$ une variété miroitée de type topologique (M, P). Une* bonne surface de découpage *est une surface strictement essentielle et H-équivariante $(F, \partial F) \hookrightarrow (\hat{M}, \partial\hat{M})$, dont la projection F dans $M = |\hat{M}/H|$ est une surface proprement plongée, qui est soit un système complet de disques méridiens si M est un corps à anses, soit une surface connexe différente d'un disque et ayant du bord si $\partial M \neq \emptyset$.*

Les deux lemmes suivants constituent la première étape pour se ramener au théorème de recollement.

5.8. Lemme

Soit $(\hat{M}, H)$ une variété miroitée de type topologique (M, P). Si M est une variété hakenienne, il existe une bonne surface de découpage dans $(\hat{M}, H)$.

Démonstration du lemme 5.8 : Soit $\mathcal{K} = \hat{M}/H$ le kaleidoscope associé à la variété miroitée $(\hat{M}, H)$. En utilisant le fait que M est hakenienne, la preuve consiste à trouver un sous-kaléidoscope $\mathcal{F}$ de dimension 2, strictement essentiel (au sens des orbifolds) dans $\mathcal{K}$, et ayant les propriétés suivantes :

1) l'espace sous-jacent de $\mathcal{F}$ est soit un système complet de disques méridiens si M est un corps à anses, soit une surface connexe différente d'un disque et ayant du bord si M a du bord.

2) le nombre de sommets de $\partial\mathcal{F}$ (i.e. de points dont le groupe d'isotropie est de rang 4) est minimal parmi tous les sous-kaléidoscopes de dimension 2 strictement essentiels dans $\mathcal{K}$ et vérifiant la propriété 1).

Le relevé $\hat{F}$ dans $\hat{M}$ de $\mathcal{F}$ est une bonne surface de découpage. Pour plus de détails nous renvoyons à [Ota2, Lemme 7.3]. □

Soient (M, P) une variété apprêtée et $(F, \partial F)$ une surface proprement plongée dans $(M, \partial M)$. Dans ce qui suit, (M_F, P_F) désignera la variété apprêtée obtenue en découpant M et P le long de F, et en oubliant les composantes obtenues en découpant P qui sont des disques. Nous utiliserons des notations analogues lorsque nous découperons une variété miroitée (respectivement un kaléidoscope) le long d'une surface équivariante (respectivement d'un sous-kaléidoscope de dimension 2).

5.9. Lemme

Soient $(\hat{M}, H)$ une variété miroitée atoroïdale de type topologique (M, P) et $\hat{F} \subset \hat{M}$ une bonne surface de découpage. Il existe une variété miroitée atoroïdale $(\hat{M}', H')$ de type topologique (M_F, P_F) et un sous-groupe $K' \subset H'$ telle que la variété miroitée $(\hat{M}', K')$ soit de type topologique $(\hat{M}_{\hat{F}}, \partial\hat{M}_{\hat{F}})$.

Démonstration du lemme 5.9 : Soit $\mathcal{F} = \hat{F}/H$ le sous-kaléidoscope strictement essentiel dans $\mathcal{K} = \hat{M}/H$. En découpant $\mathcal{K}$ le long de $\mathcal{F}$, on obtient un kaléidoscope $\mathcal{K}_{\mathcal{F}}$ dont le bord contient deux copies $\mathcal{F}^+$ et $\mathcal{F}^-$ de $\mathcal{F}$. La preuve consiste alors à ciseler $\mathcal{F}^+$ et $\mathcal{F}^-$ de façon à obtenir un nouveau kaléidoscope $\mathcal{K}'$, de bord $\{\partial\mathcal{K}\}_{\mathcal{F}}$ et d'espace sous-jacent $|\mathcal{K}'| = M_F$. Il s'agit en quelque sorte d'une version relative de la proposition 5.4. Pour plus de détails nous renvoyons à [Ota2, Lemme 7.4]. La variété miroitée $(\hat{M}', H')$, associée au kaléidoscope $\mathcal{K}'$, a pour type topologique (M_F, P_F).

Soit $K' \subset H'$ le sous-groupe engendré par les symétries miroirs correspondant aux facettes de $\mathcal{F}^+$ et $\mathcal{F}^-$. La variété miroitée $(\hat{M}', K')$ a pour type topologique $(\hat{M}_{\hat{F}}, \partial\hat{M}_{\hat{F}})$. □

Supposons par récurrence que $\mathcal{HM}(M, P) = \mathcal{AM}(M, P)$ lorsque M est une variété hakenienne, qui n'est pas virtuellement fibrée sur le cercle, et que la variété apprêtée (M, P) a une complexité $c(M, P) < (\ell, g)$, avec $\ell \geq 1$ ou $\ell = 0$ et $g \geq 1$.

Les lemmes 5.8 et 5.9 permettent de montrer :

5.10. Proposition

Soit $(\hat{M}, H)$ une variété miroitée atoroïdale de type topologique (M, P), tel que M n'est pas virtuellement fibrée sur le cercle et que $c(M, P) = (\ell, g)$. Soit $\hat{F} \subset \hat{M}$ une bonne surface de découpage. La variété apprêtée $(\hat{M}_{\hat{F}}, \partial\hat{M}_{\hat{F}})$, obtenue en découpant $(\hat{M}, \partial\hat{M})$ le long de $\hat{F}$, a les propriétés suivantes :

1) *$\partial_0\hat{M}_{\hat{F}}$ est strictement incompressible* ;

2) *$(\hat{M}_{\hat{F}}, \partial\hat{M}_{\hat{F}})$ n'est pas un fibré en intervalles* ;

3) *$(\hat{M}_{\hat{F}}, \partial\hat{M}_{\hat{F}})$ est hyperbolique.*

Démonstration de la proposition 5.10 : La définition d'une bonne surface de découpage entraîne la propriété 1) et le fait que $c(M_F, P_F) < (\ell, g)$. La propriété 2) découle du fait que M n'est pas virtuellement fibrée sur le cercle.

Soit $(\hat{M}', H')$ la variété miroitée de type topologique (M_F, P_F), donnée par le lemme 5.9. Puisque $c(M_F, P_F) < (\ell, g)$, l'hypothèse de récurrence implique que $\hat{M}'$ admet une structure hyperbolique H-invariante. Cette structure hyperbolique est aussi K'-invariante, pour le sous-groupe $K' \subset H'$, donné par la lemme 5.9. En particulier la variété miroitée $(\hat{M}', K')$ appartient à $\mathcal{HM}(\hat{M}_{\hat{F}}, \partial\hat{M}_{\hat{F}})$, qui est donc non vide. Le lemme 5.6 montre alors que la variété apprêtée $(\hat{M}_{\hat{F}}, \partial\hat{M}_{\hat{F}})$ est hyperbolique. □

La proposition 5.10 et le théorème de recollement 2.1 ont pour conséquence immédiate :

5.11. Corollaire

Soit $(\hat{M}, H)$ une variété miroitée atoroïdale de type topologique (M, P), telle que M est une variété hakenienne qui n'est pas virtuellement fibrée sur le cercle. Si (M, P) a une complexité $c(M, P) = (\ell, g)$, $int(\hat{M})$ admet une structure hyperbolique complète de volume fini. □

Le théorème de rigidité de Mostow [Mos] et le théorème de Waldhausen [Wa] impliquent que le groupe H est isotope à un groupe d'isométries de $\hat{M}$. Comme H est un groupe abélien, engendré par des involutions, l'existence d'une hiérarchie H-équivariante pour $\hat{M}$ permet, comme dans [To], de montrer que H est en fait conjugué à un groupe d'isométries de $\hat{M}$. D'où la variété miroitée $(\hat{M}, H)$ appartient à $\mathcal{HM}(M, P)$. Ceci achève la démonstration du théorème d'hyperbolisation des variétés miroitées dans le cas non fibré.

L'argument donné revient à établir une version du théorème de Waldhausen [Wa] pour les kaléidoscopes (cf. [Ka, § 7]). On peut aussi, comme dans [Ota2, § 8], éviter le recours aux théorèmes de Mostow et de Waldhausen en établissant une version équivariante du théorème de recollement 2.1. □

BIBLIOGRAPHIE

[Ah] L. Ahlfors - *Finitely generated Kleinian groups*, Amer. J. Math. **86** (1964), 413-429.

[AB] L. Ahlfors, L. Bers - *Riemann's mapping theorem for variable metrics*, Ann. of Math. **72** (1960), 385-404.

[And1] E.M. Andreev - *On convex polyhedra in Lobačeskiĭ space*, Mat. USSR Sbornik **10** (1970), 413-440.

[And2] E.M. Andreev - *On convex polyhedra of finite volume in Lobačeskiĭ space*, Mat. USSR Sbornik **12** (1971), 255-259.

[BaM] H. Bass, J.W. Morgan - *The Smith Conjecture*, Academic Press, 1984.

[BeF] M. Bestvina, M. Feighn - *Stable actions of groups on real trees*, Invent. Math. **121** (1995), 287-321.

[Ber] L. Bers - *On boundaries of Teichmüller spaces and Kleinian groups*, Ann. of Math. **91** (197O), 570-600.

[Bes] M. Bestvina - *Degenerations of hyperbolic space*, Duke Math. J. **56** (1988), 143-161.

[Bon] F. Bonahon - *Bouts des variétés hyperboliques de dimension* 3, Ann. of Math. **124** (1986), 71-158.

[BoP] M. Boileau, J. Porti - *Geometrization of 3-orbifolds of cyclic type*, Preprint, 1998.

[BS] F. Bonahon, L. C. Siebenmann - *The characteristic tori splitting of irreducible compact three-orbifolds*, Math. Ann. **278** (1987), 441-479.

[Ca] R. Canary - *A covering theorem for hyperbolic 3-manifolds and its applications*, Topology **35** (1996), 751-778.

[CEG] R. D. Canary, D. Epstein and P. Green - *Notes on notes of Thurston*, Analytical and Geometric Aspects of Hyperbolic Space (D.B.A. Epstein, ed.), London Math. Soc. Lecture Notes Ser., vol. **111**, Cambridge Univ. Press, Cambridge, 1987, 3-92.

[CJ] A. Casson, D. Jungreis - *Convergence group and Seifert fibered 3-manifolds*, Invent. Math. **118** (1994), 441-456.

[CHK] D. Cooper, C. Hodgson, S. Kerchkoff - The Orbifold Theorem: Working Notes, 1998.

[CS] M. Culler, P.B. Shalen - *Varieties of group representations and splittings of 3-manifolds*, Ann. of Math. **117** (1983), 109-146.

[DD] W. Dicks, M.J. Dunwoody - *Groups acting on graphs*, Cambridge University Press, Cambridge, 1989.

[EM] D. Epstein, A. Marden - *Convex hulls in hyperbolic space, a theorem of Sullivan, and measured pleated surfaces*, Analytical and Geometric Aspects of Hyperbolic Space (D. B. A. Epstein, ed.), London Math. Soc. Lecture Notes Ser., vol. **111**, Cambridge Univ. Press, Cambridge, 1987, 113-253.

[FLP] A. Fathi, F. Laudenbach, et V. Poenaru - *Travaux de Thurston sur les surfaces*, Astérisque **66-67** (1979).

[Ga1] D. Gabai - *Convergence groups are Fuchsian groups*, Ann. of Math. **136** (1992), 447-510.

[Ga2] D. Gabai - *Homotopy hyperbolic 3-manifolds are virtally hyperbolic*, J.A.M.S. **7** (1994), 193-198.

[Ga3] D. Gabai - *On the geometric and topological rigidity of hyperbolic 3-manifolds*, J.A.M.S. **10** (1997), 37-74.

[GMT] D. Gabai, R. Meyerhoff, N. Thurston - *Homotopy hyperbolic 3-manifolds are hyperbolic*, Preprint, 1996.

[Gro] M. Gromov - *Hyperbolic manifolds according to Thurston and Jørgensen*, Sém. Bourbaki, exp. n° 546 (novembre 1979), Lecture Notes in Math. **842** (1981), Springer Verlag, Berlin, 40-53.

[GLP] M. Gromov, J. Lafontaine, P. Pansu - *Structures métriques pour les variétés riemanniennes*, Cedic/Fernand Nathan, Paris, 1981.

[Gui] V. Guirardel - *Approximations of stable actions on $\mathbb{R}$-trees*, Comment. Math. Helv. **73** (1998), 89-121.

[JS] W.H. Jaco, P. B. Shalen - *Seifert fibred spaces in 3-manifolds*, Mem. Amer. Math. Soc. **220** (1979).

[Jo] K. Johannson - *Homotopy equivalences of 3-manifolds with boundary*, Lecture Notes in Math. **761**, Springer Verlag, Berlin, 1979.

[JM] T. Jørgensen, A. Marden - *Algebraic and geometric convergence of Kleinian groups*, Math. Scand. **66** (1990), 47-72.

[Ka] M. Kapovich - *Hyperbolic Manifolds and Discrete Groups*, Univ. of Utah Lecture Notes, 1995.

[Mard] A. Marden - *The geometry of finitely generated Kleinian groups*, Ann. of Math. **99** (1974), 383-496.

[Marg] G.A. Margulis - *Discrete groups of motions of manifolds of non-positive curvature*, Proc. Int. Congr. Math., Vancouver **2** (1974), 21-34.

[Mas1] B. Maskit - *On Klein's combination theorem* III, Advances in the theory of Riemann Surfaces (L. Ahlfors et al., ed.), Ann. of Math. Studies **66**, Princeton Univ. Press (1971), 297-316.

[Mas2] B. Maskit - *Kleinian Groups*, Springer Verlag, Berlin, 1988.

[MT] K. Matsuzaki, M. Taniguchi - *Hyperbolic manifolds and Kleinian groups*, Oxford Math. Monographs, Oxford, 1998.

[McM1] C. McMullen - *Amenability, Poincaré series and holomorphic averaging*, Invent. Math. **97** (1989), 95-127.

[McM2] C. McMullen - *Iteration on Teichmüller space*, Invent. Math. **99** (1990), 425-454.

[McM3] C. McMullen - *Riemann surface and the geometrization of 3-manifolds*, Bull. Amer. Math. Soc. **27** (1992), 207-216.

[McM4] C. McMullen - *Renormalization and 3-manifolds which Fiber over the Circle*, Annals of Math. Studies, vol. 144, Princeton University Press, 1996.

[MY] W.H. Meeks, S.T. Yau - *The equivariant Dehn's Lemma and Loop Theorem*, Comment. Math. Helv. **56** (1981), 225-239.

[Me] G. Mess - *The Seifert conjecture and groups which are coarse quasi-isometric to planes*, Preprint, 1988.

[Mor1] J. Morgan - *On Thurston's uniformization theorem for three-dimensional manifolds*, The Smith Conjecture (H. Bass et J. Morgan, ed.), Academic Press: Orlando, Fl., 1984, 37-125.

[Mor2] J. Morgan - *Actions de groupes finis sur* $\mathbb{S}^3$ *: la conjecture de P.A. Smith (d'après Thurston et Meeks-Yau)*, Sém. Bourbaki, exp. n° 578 (juin 1981), Lecture Notes in Math. **901** (1981), Springer Verlag, Berlin, 277-289.

[MS1] J. Morgan, P.B. Shalen - *Valuations, trees, and degenerations of hyperbolic structures, I*, Ann. of Math. **120** (1984), 401-476.

[MS2] J. Morgan, P.B. Shalen - *Valuations, trees, and degenerations of hyperbolic structures, II: measured laminations in 3-manifolds*, Ann. of Math. **127** (1988), 403-465.

[MS3] J. Morgan, P.B. Shalen - *Valuations, trees, and degenerations of hyperbolic structures, III : actions of 3-manifold groups on trees and Thurston's compactness theorem*, Ann. of Math. **120** (1984), 467-519.

[Mos] G.D. Mostow - *Strong rigidity of locally symmetric spaces*, Ann. of Math. Stud., vol. 78, Princeton University Press, Princeton, NJ, 1973.

[Oh] K. Ohshika - *Geometric behavior of Klenian groups on boundaries of deformation spaces*, Quart. J. Math. Oxford **43** (1992), 97-111.

[Ota1] J.-P. Otal - *Le théorème d'hyperbolisation pour les variétés fibrées de dimension* 3, Astérisque **235** (1996).

[Ota2] J.-P. Otal - *Thurston's hyperbolization of Haken manifolds*, Preprint, 1997.

[Pau1] F. Paulin - *Topologie de Gromov équivariante, structures hyperboliques et arbres réels*, Invent. Math. **94** (1988), 53-80.

[Pau2] F. Paulin - *Actions de groupes sur les arbres*, Sém. Bourbaki, exp. n° 808 (novembre 1995), Lecture Notes in Math. **241** (1997), Springer Verlag, Berlin, 97-137.

[Ril] R. Riley - *A quadratic parabolic group*, Math. Proc. Camb. Phil. Soc. **77** (1975), 281-288.

[Riv] I. Rivin - *A characterization of ideal polyhedra in hyperbolic* 3*-space*, Ann. of Math. **143** (1996), 51-70.

[RH] I. Rivin, C. Hodgson - *A characterization of compact convex polyhedra in hyperbolic* 3*-space*, Invent. Math. **111** (1993), 77-111.

[Sco1] P. Scott - *A new proof of the annulus and torus theorems*, Amer. J. Math. **102** (1980), 241-277.

[Sco2] P. Scott - *There are no fake Seifert fibre spaces with infinite* π_1, Ann. of Math. **117** (1983), 35-70.

[Sco3] P. Scott - *The geometries of* 3*-manifolds*, London Math. Soc. **15** (1983), 401-487.

[Sko] R. Skora - *Splittings of surfaces*, J.A.M.S. **9** (1996), 605-616.

[Sta] J. Stallings - *On fibering certain* 3*-manifolds*, Topology of 3-manifolds and related topics, Prentice Hall (1962), 95-100.

[Su1] D. Sullivan - *Travaux de Thurston sur les groupes quasi-fuchsiens et sur les variétés hyperboliques de dimension* 3 *fibrées sur* S^1, Sém. Bourbaki, exp. n° 554 (février 1980), Lecture Notes in Math. **842** (1982), Springer Verlag, Berlin, 196-214.

[Su2] D. Sullivan - *On finiteness theorem for cusps*, Acta Math. **147** (1981), 289-299.

[Th0] W.P. Thurston - *The geometry and topology of* 3*-manifolds*, Princeton Math. Dept., 1979.

[Th1] W.P. Thurston - *Hyperbolic structures on* 3*-manifolds: overall logic*, Notes of Summer Workshop at Bowdoin (1980).

[Th2] W.P. Thurston - *Three dimensional manifolds, Kleinian groups and hyperbolic geometry*, Bull. Amer. Math. Soc. **6** (1982), 357-381.

[Th3] W.P. Thurston - *Hyperbolic Structures on 3-manifolds, I: Deformations of acylindrical manifolds*, Annals of Math. **124** (1986), 203-246.

[Th4] W.P. Thurston - *Hyperbolic Structures on 3-manifolds, II: Surface groups and manifolds which fiber over* S^1, Preprint, 1986.

[Th5] W.P. Thurston - *Hyperbolic Structures on 3-manifolds, III: Deformations of 3-manifolds with incompressible boundary*, Preprint, 1986.

[Th6] W.P. Thurston - *Three-manifolds with symmetry*, Preprint, 1982.

[Th7] W.P. Thurston - *On the geometry and dynamics of diffeomorphisms of surfaces*, Bull. Amer. Math. Soc. **19** (1988), 417-432.

[Th8] W.P. Thurston - *Three-dimensional geometry and topology*, vol. 35 (S. Levy, ed.), Princeton University Press, 1997.

[To] J.L. Tollefson - *Involutions of sufficiently large* 3*-manifolds*, Topology **20** (1981), 323-352.

[Tu] P. Tukia - *Homeomorphic conjugates of Fuchsian groups*, J. Reine Angew. Math. **391** (1988), 1-54.

[Wa] F. Waldhausen - *On irreducible* 3*-manifolds which are sufficiently large*, Ann. of Math. **87** (1968), 56-88.

[WS] C. Weber, H. Seifert - *Die beiden Dodekaederräume*, Math. Z. **37** (1933), 237-253.

Michel BOILEAU

Laboratoire Émile Picard
UMR 5580 du CNRS
Université Paul Sabatier
118, route de Narbonne
F-31062 TOULOUSE Cedex 04
E-mail : `boileau@picard.ups-tlse.fr`

Astérisque

NICOLAS BURQ

Formules de trace, résonances et quasi-modes

Astérisque, tome 266 (2000), Séminaire Bourbaki,
exp. n° 856, p. 175-190

<http://www.numdam.org/item?id=SB_1998-1999__41__175_0>

© Société mathématique de France, 2000, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (http://smf4.emath.fr/Publications/Asterisque/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Séminaire BOURBAKI
51ème année, 1998-99, n° 856

Mars 1999

FORMULES DE TRACE, RÉSONANCES ET QUASI-MODES
[d'après Sjöstrand-Zworski, Stefanov-Vodev et al.]

par **Nicolas BURQ**

1. INTRODUCTION

On se propose dans cet exposé de présenter quelques résultats récents mettant en évidence l'existence de *résonances*. Rappelons par exemple qu'à l'extérieur d'un obstacle compact $\Theta \subset \mathbb{R}^n$, le laplacien avec conditions de Dirichlet Δ_D, sur l'espace $L^2(\Theta^c)$, de domaine l'espace de Sobolev $H^2 \cap H_0^1(\Theta^c)$ est autoadjoint et que sa résolvante $(\mathcal{P} + \mu^2)^{-1}$ est un opérateur holomorphe dans l'ensemble $\mathbf{Im}\mu > 0$. On peut montrer que cette résolvante (agissant de L^2_{comp} dans L^2_{loc}) admet un prolongement méromorphe à $\mathbb{C}$ si n est impair et au revêtement simplement connexe $\widetilde{\mathbb{C}}$ de $\mathbb{C}^*$ si n est pair. Les résonances sont les pôles de ce prolongement (nous préciserons cette définition dans la section 2). Elles peuvent être interprétées de plusieurs façons. Outre leur définition, ce sont aussi les points pour lesquels il existe une solution sortante (vérifiant les conditions de radiation de Sommerfeld à l'infini, $|\partial_r u + i\mu u| \leq C|e^{-i\mu r}|/r^{1+n/2}$) non triviale de l'équation

$$(\Delta + \mu^2)\, u = 0 \; u\,|_{\partial\Omega} = 0 \; ; \tag{1.1}$$

ce sont également (en dimension impaire d'espace) les pôles de la matrice de diffusion (scattering) divisés par le complexe i et les valeurs propres du générateur infinitésimal du semi-groupe de Lax et Phillips. Ce sont enfin, nous le verrons à la section 2, les valeurs propres de certains opérateurs (non autoadjoints) construits à partir de $-\Delta_D$ par *distorsion analytique*. Les résonances peuvent permettre (en dimension d'espace impaire) de donner un développement asymptotique en grand temps pour les solutions de l'équation des ondes dans Ω, du type

$$U(t)\, f = \sum_{\mathbf{Im}\mu < C} e^{it\mu} \Pi_\mu(f) + \mathcal{O}\left(e^{-(C-\varepsilon)t}\right),$$

où Π_μ est le projecteur spectral sur la "fonction propre" associée au pôle μ vérifiant (1.1) et $U(t)\, f$ la solution de l'équation des ondes avec conditions au bord de Dirichlet et de données initiales $u\,|_{t=0} = 0, \partial_t u\,|_{t=0} = f$.

Pour l'opérateur $-\Delta_D$, les *quasi-modes* sont des suites $(u_n, \mu_n)_{n\in\mathbb{N}}$, où $\mu_n \in \mathbb{R}$, $u_n \in H^2 \cap H_0^1(\Theta^c)$ est une fonction à support compact vérifiant $\|u_n\|_{L^2} = 1$ et $\mu_n \overset{n\to+\infty}{\longrightarrow} +\infty$, vérifiant pour tout $N > 0$:

$$\left(\Delta + \mu_n^2\right) u_n = O\left(\mu_n^{-N}\right) \; u_n \,|_{\partial\Omega} = 0.$$

De tels quasi-modes ont été construits (voir les résultats de V.M. Babich-V.S. Buldyrev [2], J. Ralston [10], V.F. Lazutkin [6], G. Popov [9]) depuis les années 70 dans certaines géométries favorables (présence de trajectoires captives de type elliptique pour le flot du billard dans Θ^c par exemple).

Nous nous sommes intéressé dans cet exposé à des résultats montrant la présence de résonances, dans un cadre bien plus général que celui du laplacien de Dirichlet. Ces résultats sont de deux types très différents : d'une part, nous présenterons des résultats de J. Sjöstrand-M. Zworski [17, 19, 14, 15] utilisant de nouvelles formules de traces valides en dimensions quelconques d'espace (et plus seulement en dimensions impaires) permettant d'obtenir des bornes inférieures pour les fonctions de comptage des résonances (et en particulier de montrer la présence de résonances), et d'autre part des résultats de P. Stefanov-G. Vodev [21, 20], S.U. Tang-M. Zworski [22] montrant que les quasi-modes sont en fait asymptotiquement proches de résonances.

Le plan de l'exposé est le suivant : dans une première partie nous présentons le cadre mathématique de la "boîte noire" de J. Sjöstrand-M. Zworski permettant de définir les résonances pour un opérateur semi-classique analytiquement dilatable à l'infini. Même si elle ne contient pas tous les cas pour lesquels il est possible de définir les résonances, cette approche fournit le cadre conceptuel le plus utilisable à l'heure actuelle pour un analyste. Dans une deuxième partie nous présentons une nouvelle formule de trace due à J. Sjöstrand et nous montrons comment cette formule permet dans l'esprit de J. Sjöstrand-M. Zworski [17] de donner une borne inférieure (non triviale) pour la fonction de comptage des résonances pour l'opérateur de Schrödinger, $-h^2\Delta+V$, sous des hypothèses très faibles sur le potentiel (les hypothèses sont par exemple vérifiées par tout potentiel réel régulier et à support compact). Enfin dans une troisième partie nous présentons les résultats de P. Stefanov-G. Vodev, S.U. Tang-M. Zworski mettant en évidence les liens entre quasi-modes et résonances.

Cet exposé ne prétend pas être exhaustif, ni en ce qui concerne les travaux des auteurs cités dont nous ne présentons qu'une faible part (nous ne parlerons en particulier ni des bornes supérieures pour les fonctions de comptage des résonances ni de résultats récents de J. Sjöstrand-M. Zworski [19] sur la distribution asymptotique des résonances à l'extérieur d'un obstacle convexe), ni en ce qui concerne de nombreux travaux d'autres auteurs sur le sujet.

2. LES RÉSONANCES

2.1. La *boîte noire*

On considère un espace de Hilbert complexe $\mathcal{H}$ et une famille d'opérateurs $\mathcal{P}$, autoadjoints et non bornés sur $\mathcal{H}$, de domaines $\mathcal{D}$, dépendant d'un paramètre $h \in]0, h_0]$. On munit $H^2(\mathbb{R}^n \setminus B(0,R_0))$ de la norme $\|u\|_{H^2}^2 = \|(1-h^2\Delta)u\|_{L^2}$ et $\mathcal{D}$ de la norme $\|u\|_{\mathcal{D}} = \|(i+\mathcal{P})^{-1}u\|_{L^2}$.

Nous ferons les hypothèses suivantes sur $\mathcal{H}$ et $\mathcal{P}$.

- **H1** L'espace $\mathcal{H}$ admet une décomposition orthogonale

$$\mathcal{H} = \mathcal{H}_{R_0} \oplus L^2(\mathbb{R}^n \setminus B(0,R_0)),$$
$$u \in \mathcal{H} = 1_{B(0,R_0)}u + 1_{\mathbb{R}^n \setminus B(0,R_0)}u$$

et on suppose que

$$1_{\mathbb{R}^n \setminus B(0,R_0)}\mathcal{D} = H^2(\mathbb{R}^n \setminus B(0,R_0)),$$

uniformément par rapport à h, c'est-à-dire que l'application

$$1_{\mathbb{R}^n \setminus B(0,R_0)} : \mathcal{D} \to H^2(\mathbb{R}^n \setminus B(0,R_0))$$

est uniformément bornée par rapport à h et à un inverse à droite qui est aussi uniformément borné par rapport à h.

- **H2** Si $u \in H^2(\mathbb{R}^n \setminus B(0,R_0))$ est nulle au voisinage de $\overline{B(0,R_0)}$, alors u (ou plutôt son prolongement par 0 en dehors de $\mathbb{R}^n \setminus B(0,R_0)$) appartient à $\mathcal{D}$.
- **H3** L'opérateur $1_{B(0,R_0)}(\mathcal{P}+i)^{-1}$ est compact.
- **H4** On suppose que $(\mathcal{P}u)|_{\mathbb{R}^n \setminus \overline{B(0,R_0)}}$ est de la forme $Q(u)$, avec

$$Qu = \sum_{|\alpha|=2} a_\alpha(x)(hD_x)^\alpha u + \sum_{|\alpha|<2} a_\alpha(x,h)(hD_x)^\alpha u, \tag{2.1}$$

où Q est formellement autoadjoint, les fonctions a_α sont uniformément bornées par rapport à h ainsi que toutes leurs dérivées et vérifient

$$\begin{gathered}\exists c > 0; \forall x,\ \textstyle\sum_{|\alpha|=2} a_\alpha \xi^\alpha \geq c|\xi|^2, \\ \textstyle\sum_{|\alpha|\leq 2} a_\alpha \xi^\alpha \overset{x\to+\infty}{\longrightarrow} |\xi|^2 \text{ uniformément par rapport à } h.\end{gathered} \tag{2.2}$$

- **H5** Soient $R > 2R_0$ et $\mathcal{P}^\sharp$ l'opérateur sur $\mathcal{H}^\sharp = \mathcal{H}_{R_0} \oplus L^2(\mathbb{R}^n/R\mathbb{Z}^n \setminus B(0,R_0))$ qui coïncide avec $\mathcal{P}$ sur la cellule $[-R,R]^n$ avec des conditions périodiques au bord. L'opérateur $\mathcal{P}^\sharp$ a alors un spectre discret et on note $N^\sharp(\mathcal{P}^\sharp, I)$ le nombre de ses valeurs propres dans un intervalle I. On suppose qu'on a une formule de type *formule de Weyl* pour la fonction de comptage, $N^\sharp$, des valeurs propres de $\mathcal{P}^\sharp$:

$$N^\sharp(\mathcal{P}^\sharp, [-\lambda,\lambda]) = \mathcal{O}(\lambda/h^2)^{n^\sharp}; \lambda \geq 1.$$

On peut vérifier que cette hypothèse ne dépend pas du choix de R.

Les deux exemples fondamentaux d'opérateurs vérifiant les hypothèses précédentes sont les suivants :

L'opérateur de Schrödinger. On considère $\mathcal{P} = -h^2\Delta + V(x,h)$ où le potentiel V est réel et vérifie $\lim_{x\to+\infty} V(x,h) = 0$ (uniformément par rapport à h), sur l'espace de Hilbert $\mathcal{H} = L^2(\mathbb{R}^n)$ et de domaine $\mathcal{D} = H^2(\mathbb{R}^n)$.

Le problème extérieur. On considère $\Theta \subset \mathbb{R}^n$ un domaine borné régulier, $\Omega = \mathbb{R}^n \setminus \overline{\Theta}$ et $\mathcal{P} = -h^2\Delta_D$ (respectivement $\mathcal{P} = -h^2\Delta_N$) le laplacien de Dirichlet (respectivement de Neumann) sur l'espace de Hilbert $\mathcal{H} = L^2(\Omega)$ de domaine $\mathcal{D}_D = H^2(\Omega) \cap H_0^1(\Omega)$ (respectivement $H^2(\Omega)$).

2.2. **Distorsion analytique**

Dans cette partie on définit les résonances comme étant les valeurs caractéristiques de certains opérateurs non autoadjoints, obtenus à partir de P par distorsion analytique. L'introduction de la distorsion analytique pour l'étude des résonances remonte aux travaux de J. Aguilar-J.M. Combes [1].

- **H6** On suppose qu'il existe un angle $\theta_0 \in]0,\pi[$ et $\varepsilon > 0$, $R_1 > R_0$ tels que les coefficients $a_\alpha(x,h)$; $|\alpha| \leq 2$ admettent un prolongement holomorphe en x à l'ensemble

$$\Omega_{R_0,\theta_0,\varepsilon} = \{r\omega; \omega \in \mathbb{C}^n; \operatorname{dist}(\omega, \mathbb{S}^{n-1}) < \varepsilon, r \in e^{i[0,\theta_0[}]R_1, +\infty[\}.$$

On suppose aussi que les relations (2.1) et (2.2) restent vérifiées dans cet ensemble. Cette hypothèse supplémentaire est évidemment vérifiée pour le problème extérieur et pour l'opérateur de Schrödinger si par exemple le potentiel est à support compact.

Les constantes $\varepsilon_0 > 0, R_1 > R_0$ étant fixées, on peut construire une famille de déformations

$$[0,\theta_0] \times [0,+\infty[\ni (\theta,t) \mapsto f_\theta(t) \in \mathbb{C},$$

vérifiant

(1) $f_\theta(t) = t$ pour $t \leq R_1$,

(2) $0 \leq \arg(f_\theta(t)) \leq \theta$, $\partial_t f_\theta \neq 0$,

(3) $\arg(f_\theta(t)) \leq \arg\partial_t(f_\theta(t)) \leq \arg(f_\theta(t)) + \varepsilon_0$,

(4) $f_\theta(t) = e^{i\theta}t$ pour $t \geq T_0$, T_0 dépendant de R_1 et ε_0.

On considère alors l'application

$$\kappa_\theta : \mathbb{R}^n \ni x = |x|\omega \mapsto f_\theta(|x|)\,\omega \in \mathbb{C}^n$$

et on note Γ_θ son image qui est une variété totalement réelle ($T\Gamma_\theta \cap iT\Gamma_\theta = 0$) qui coïncide avec $\mathbb{R}^n$ près de $B(0,R_1)$. On note aussi

$$\mathcal{H}_\theta = \mathcal{H}_{R_0} \oplus L^2(\Gamma_\theta \setminus B(0,R_0)).$$

La fonction κ_θ permet d'identifier $\mathcal{H}$ et $\mathcal{H}_0$. De la même façon, on définit $\mathcal{D}_\theta = \kappa_{\theta *}(\mathcal{D})$. Soit $\chi \in C_0^\infty(B(0,R_1))$ égale à 1 au voisinage de $B(0,R_0)$. On définit alors un opérateur $\mathcal{P}_\theta$ non borné (et non autoadjoint) de domaine $\mathcal{D}_\theta$ par

$$\mathcal{P}_\theta u = \mathcal{P}(\chi u) + Q\,|_{\Gamma_\theta}((1-\chi)u);\ 0 \le \theta \le \theta_0.$$

Le symbole principal de l'opérateur $Q\,|_{\Gamma_\theta}$ vaut $q_\theta(z,\zeta) = q_0\left(\left(\kappa_\theta^{-1}(z),\ {}^tT\kappa_\theta(\zeta)\right)\right)$. Si on identifie Γ_θ et Γ, on obtient par exemple pour $-\Delta$ et son symbole principal, D en coordonnées polaires :

$$-\Delta_{\Gamma_\theta} = \left(\frac{1}{f'(t)}D_t\right)^2 - \frac{n-1}{f(t)f'(t)}iD_t + (f(t))^{-2}D_\omega^2,$$
$$D_\theta = \frac{\tau^2}{(f'(t))^2} + \frac{|\eta|^2}{f(t)^2}.$$

Les lemmes suivants permettent de définir les résonances dans ce cadre : le premier résultat garantit que le spectre de $\mathcal{P}_\theta$ est en dehors de $e^{-2i\theta}[0,+\infty[$ formé de valeurs propres :

LEMME 2.1 (J. Sjöstrand-M. Zworski [16]). — *Soit $z \in \mathbb{C}^*$, tel que $\arg z \neq -2\theta$. Alors $(\mathcal{P}_\theta - z) : \mathcal{D}_\theta \mapsto \theta$ est un opérateur de Fredholm d'indice 0. En particulier un point $z \in \mathbb{C} \setminus e^{-2i\theta}[0,+\infty[$ appartient au spectre de $\mathcal{P}_\theta$ si et seulement si $\dim Ker(\mathcal{P}_\theta - z) \neq 0$*

et le deuxième résultat décrit l'invariance du spectre par rapport à θ et au choix de la déformation f_θ :

LEMME 2.2 (J. Sjöstrand-M. Zworski [16]). — *Soient $0 \le \theta_1 \le \theta_2 \le \theta_0$ et $z \in \mathbb{C} \setminus e^{-2i[\theta_1,\theta_2]}[0,+\infty[$. Alors*

$$\forall k \in \mathbb{N}, \dim \mathrm{Ker}\,(\mathcal{P}_{\theta_1} - z)^k = \dim \mathrm{Ker}\,(\mathcal{P}_{\theta_2} - z)^k.$$

De plus, si on choisit une deuxième famille de déformation, $\widetilde{f}_\theta$ vérifiant les hypothèses **(1)**–**(4)** *et si on note $\widetilde{\mathcal{P}}_\theta$ la famille d'opérateurs associée, alors*

$$\forall k \in \mathbb{N}, \dim \mathrm{Ker}\big(\widetilde{\mathcal{P}}_{\theta_1} - z\big)^k = \dim \mathrm{Ker}\big(\mathcal{P}_{\theta_1} - z\big)^k.$$

DÉFINITION 2.3. — *On appelle* résonances *de l'opérateur $\mathcal{P}$ incluses dans $e^{-2i[0,\theta[}]0,+\infty[$ les valeurs propres de l'opérateur $\mathcal{P}_\theta$ incluses dans $e^{-2i[0,\theta[}]0,+\infty[$. On notera* Rés$(\mathcal{P})$ *l'ensemble des résonances de $\mathcal{P}$.*

D'après le lemme 2.2, cette définition ne dépend ni du choix de θ ni de celui de la famille de déformations f_θ. D'après la théorie de Fredholm analytique, si $z_0 \in e^{-2i[0,\theta[}]0,+\infty[$ est une résonance, alors l'opérateur

$$\pi_{\theta,z_0} = \frac{1}{2\pi i}\int_{|z-z_0|=\eta<<1} (z - \mathcal{P}_\theta)^{-1}\,dz$$

est une projection de rang fini, son image F_{θ,z_0} est, pour $k \in \mathbb{N}$ assez grand, le noyau de $(\mathcal{P}_\theta - z_0)^k$ et sa dimension ne dépend donc ni du choix de θ ni de celui de f_θ.

DÉFINITION 2.4. — *Si* $z_0 \in e^{-2i[0,\theta[}]0,+\infty[$ *est une résonance alors on définit sa* multiplicité mult (z_0) *par*

$$\text{mult}\,(z_0) = \dim F_{\theta,z_0}.$$

REMARQUE 2.5. — La définition des résonances que nous avons donnée coïncide avec les autres définitions possibles, en particulier, on aurait pu montrer que la résolvante $(z-\mathcal{P})^{-1}$ définie (et holomorphe) de $\mathcal{H}_{\text{comp}}$ dans $\mathcal{H}_{\text{loc}}$ pour $\mathbf{Im}z > 0$ peut être prolongée méromorphiquement à travers $]0,+\infty[$ et que les pôles de ce prolongement sont les résonances. Plus précisément, on a le résultat suivant :

LEMME 2.6. — *Soit* $\chi \in C_0^\infty(\mathbb{R}^n)$ *égale à* 1 *au voisinage de* $B(0,R_0)$. *Alors la résolvante tronquée,* $\chi(\mathcal{P}-z)^{-1}\chi$, *qui est holomorphe pour* $\mathbf{Im}z > 0$, *admet un prolongement méromorphe à l'ensemble* $z \in e^{-2i]\pi,\theta[}]0,+\infty[$, *les pôles de ce prolongement sont les résonances situées dans* $e^{-2i]0,\theta[}]0,+\infty[$ *et leur multiplicité est la dimension de l'image du projecteur spectral tronqué correspondant :*

$$\chi\pi_{z_0}\chi = \frac{1}{2\pi i}\int_{|z-z_0|=\eta<<1} \chi(z-\mathcal{P})^{-1}\chi dz.$$

REMARQUE 2.7. — La définition des résonances que nous avons esquissée dans l'introduction correspond donc à prendre $\mu = h^{-1}\sqrt{\lambda}$ où λ est une résonance suivant la définition 2.3 et $\sqrt{\lambda}$ la racine est la détermination principale de la racine (positive sur $\mathbb{R}^{+*}$).

3. FORMULES DE TRACE ET RÉSONANCES

3.1. Formules de trace

Dans le cas de la diffusion par un obstacle, Θ, ou par un potentiel à support compact, notons $U(t) = \cos\left(t\sqrt{\mathcal{P}}\right)$, $U_0(t) = \cos\left(t\sqrt{\mathcal{P}_0}\right)$, $U_0(t) = 1_{\Theta^c}\cos\left(t\sqrt{\mathcal{P}_0}\right)1_{\Theta^c}$ dans le cas d'un obstacle, $\mathcal{P}_0 = -\Delta$ dans $\mathbb{R}^n$ (ici on fixe $h=1$). La formule de Poisson suivante a été démontrée par P. Lax-R. Phillips [5], puis généralisée successivement par C. Bardos-J.C. Guillot-J. Ralston [3], R. Melrose, [7, 8] et J. Sjöstrand-M. Zworski [18] *dans le cas où la dimension d'espace est impaire* :

$$u(t) = 2\text{Tr}\,(U(t) - U_0(t)) = \sum_{\lambda\in\widetilde{\text{Rés}}(\mathcal{P})} e^{-i\lambda|t|};\ t \neq 0,$$

l'égalité ayant lieu au sens des distributions sur $\mathbb{R}^*$, et $\widetilde{\text{Rés}}$ désignant l'ensemble des résonances telles que définies dans l'introduction. Cette relation s'exploite souvent de la manière suivante : on considère $\chi \in C_0^\infty(]0,+\infty[)$ et on obtient pour $\lambda \in \mathbb{R}$

$$\text{(3.1)}\quad \text{Tr}\left(\widehat{\chi}\left(\lambda-\sqrt{\mathcal{P}}\right)+\widehat{\chi}\left(\lambda+\sqrt{\mathcal{P}}\right)-\widehat{\chi}\left(\lambda-\sqrt{\mathcal{P}_0}\right)-\widehat{\chi}\left(\lambda+\sqrt{\mathcal{P}_0}\right)\right) = \widehat{\chi u}(\lambda) = \sum_{\mu\in\widetilde{\text{Rés}}} \widehat{\chi}(\lambda-\mu).$$

La fonction

$$\widehat{\chi}\left(\lambda-\sqrt{z}\right)+\widehat{\chi}\left(\lambda+\sqrt{z}\right)$$

étant holomorphe (car les puissances impaires de $\sqrt{z}$ se compensent). On s'intéresse alors à l'asymptotique $\lambda \to +\infty$ de la relation (3.1). Posons $\lambda = h^{-1}$; on obtient

$$\widehat{\chi u}(\lambda) = \sum_{\mu \in \widetilde{\text{Rés}}} \widehat{\chi}\left(h^{-1}\left(1-h\sqrt{\mu}\right)\right) + \widehat{\chi}\left(h^{-1}\left(1+h\sqrt{\mu}\right)\right).$$

On constate que la contribution du second terme est $\mathcal{O}(h^{\infty})$ tandis que celle du premier terme donne

$$\widehat{\chi u}(\lambda) = \sum_{\mu \in \text{Rés}} \widehat{\chi}\left(h^{-1}\left(1-\mu\right)\right) + \mathcal{O}(h^{\infty}); \tag{3.2}$$

des renseignements sur les singularités de u (et donc sur la non décroissance rapide en λ de $\widehat{\chi u}$) peuvent alors permettre d'obtenir des indications sur la localisation des résonances.

Le but de cette partie est de présenter une formule de trace semi-classique locale due à J. Sjöstrand applicable en toutes dimensions d'espace pour des perturbations du Laplacien qui peuvent être à longue portée et permettant de fournir des résultats du type (3.2). Nous suivons ici la présentation de [15]. Nous noterons par la suite, si f_0 et f_1 sont deux quantités, $[f.]_0^1 = f_1 - f_0$.

On considère deux opérateurs $\mathcal{P}_{0;1}$ de domaines $\mathcal{D}_{0;1}$ vérifiant les hypothèses **H1–6**. On suppose qu'ils vérifient l'hypothèse supplémentaire :

- **H7**

$$\forall \alpha; |\alpha| \leq 2, \forall 0 < h \leq h_0,\ \forall x \in \Omega_{R_0,\theta,\varepsilon},$$
$$\left|\left[a_{;\alpha}(x,h)\right]_0^1\right| = \left|\left[a_{1,\alpha}(x,h) - a_{0,\alpha}(x,h)\right]_0^1\right| \leq C\left(1+|x|\right)^{-\widetilde{n}}; \widetilde{n} > n.$$

Le résultat suivant permet de définir la trace (on note $n^{\sharp} = \max\{n_0^{\sharp}, n_1^{\sharp}\}$) :

PROPOSITION 3.1. — *Soient $m > 2N$, $\chi \in C_0^{\infty}(\mathbb{R}^n)$, égale à 1 au voisinage de $B(0,R_0)$ et $f \in C_0^{\infty}(\mathbb{R})$. Alors les opérateurs*

$$\chi f\left(\mathcal{P}_{0;1}\right), f\left(\mathcal{P}_{0;1}\right)\chi \text{ et } \left[\left(1-\chi\right) f\left(\mathcal{P}.\right)\left(1-\chi\right)\right]_0^1$$

sont à trace et l'expression

$$Tr\Big[f\left(\mathcal{P}.\right)\Big]_0^1 \overset{def}{=} \Big[Tr\left(f\left(\mathcal{P}.\right)\chi + \chi f\left(\mathcal{P}.\right)\left(1-\chi\right)\right)\Big]_0^1 + Tr\Big[\left(1-\chi\right) f\left(\mathcal{P}.\right)\left(1-\chi\right)\Big]_0^1$$

est indépendante du choix de χ et majorée par $Ch^{-n^{\sharp}}$.

Soient $W \subset\subset \Omega$ deux ouverts relativement compacts de $e^{i]-2\theta_0,\varepsilon_0]}]0,+\infty[$. On suppose que les intersections I et J de W et Ω avec $\mathbb{R}^+$ sont des intervalles et que Ω est simplement connexe. On note Ω_- et W_- les intersections de Ω et W avec $e^{i]-2\theta_0,0]}]0,+\infty[$.

THÉORÈME 1 (Sjöstrand, 96). — *Soit $f(z,h)$ une famille de fonctions holomorphes en $z \in \Omega$ telle que $|f|_{\Omega\setminus W}| \leq 1$. Soit $\chi \in C_0^\infty(\mathbb{R})$ égale à 1 au voisinage de $\overline{I}$. Alors*

$$\mathrm{Tr}\left((\chi f)(\mathcal{P}_1) - (\chi f)(\mathcal{P}_0)\right) = \Big[\sum_{\lambda \in \mathrm{Rés}\mathcal{P}_\cdot \cap W_-} f(\lambda,h) - \sum_{\mu \in \sigma(\mathcal{P}_\cdot)\cap]-\infty,0[\cap W_-} f(\mu,h)\Big]_0^1 + \mathcal{O}\left(h^{-n\sharp}\right).$$

REMARQUE 3.2. — Si $W \cap \mathbb{R}^- = \emptyset$, le choix de $f = 1$ sur W dans la formule de trace permet, si on peut construire un opérateur $\mathcal{P}_0$ n'ayant pas de résonances dans W, de retrouver que le nombre de résonances de $\mathcal{P}_1$ dans W est majoré par $Ch^{-n\sharp}$ (ce résultat était connu antérieurement).

La démonstration de cette formule de trace est assez technique. Nous allons en donner les grandes lignes. On fixe $m \in \mathbb{N}$ assez grand pour que les opérateurs qu'on considère aient de bonnes propriétés de trace, on écrit, pour un $z_0 \notin W$ de partie imaginaire strictement positive, loin du spectre de $\mathcal{P}$, $f(z) = (z - z_0)^{-m} g(z)$ et on obtient par le calcul fonctionnel des opérateurs autoadjoints

$$\begin{aligned}(\chi f)(\mathcal{P}_\cdot) &= (\mathcal{P} - z_0)^{-m} (\chi g)(\mathcal{P}_\cdot), \\ &= -\frac{1}{2i\pi} \int_{\mathbb{R}} g(x)\chi(x)(\mathcal{P}_\cdot - z_0)^{-m}\left((x + i0 - \mathcal{P})^{-1} - (x - i0 - \mathcal{P})^{-1}\right) dx.\end{aligned}$$

Soit $\widetilde{\chi}$ une extension quasi analytique de χ ($\partial_{\bar{z}}\chi$ est nulle à l'ordre infini sur $\mathbb{R}$), à support proche de $J_+ = J \cap \mathbb{R}^+$, égale à 1 au voisinage de $I_+ = I \cap \mathbb{R}^+$. La formule de Stokes implique ($dL(z)$ étant la mesure de Lebesgue sur $\mathbb{C}$),

$$\begin{aligned}(\chi f)(\mathcal{P}_\cdot) &= -\frac{1}{\pi}\int_{\mathbb{C}} g(z)\,\partial_{\bar{z}}\widetilde{\chi}(z)(\mathcal{P}_\cdot - z_0)^{-m}(z - \mathcal{P}_\cdot)^{-1} dL(z) \\ &= -\frac{1}{\pi}\int_{\Omega_-} - \frac{1}{\pi}\int_{\Omega_+} \\ &= I_\cdot^- + I_\cdot^+.\end{aligned}$$

On choisit $\widehat{\chi}$ égale à 1 au voisinage de W_-, à $\widetilde{\chi}$ près de J_+ et quasi analytique près de $\mathbb{R}^-$. On obtient

$$I_\cdot^- = -(\widehat{\chi}1_{\mathbb{R}^-} f)(\mathcal{P}_\cdot) - \frac{1}{\pi}\int_{\mathbb{C}} g(z)\,\partial_{\bar{z}}\widehat{\chi}(z)(\mathcal{P}_\cdot - z_0)^{-m}(z - \mathcal{P}_\cdot)^{-1} dL(z).$$

On vérifie alors que la contribution du premier terme à la trace donne le terme de la formule de trace correspondant aux valeurs propres négatives des opérateurs $\mathcal{P}_{0;1}$ tandis que celle du deuxième est en $\mathcal{O}\left(h^{-n\sharp}\right)$. Il reste à étudier la contribution de I^+ à la formule de trace.

La formule de Stokes donne

$$I_\cdot^+ = -\frac{1}{\pi}\int_{\Omega_+\cap \mathbf{Im} z\leq\delta} g(z)\,\partial_{\bar z}\widetilde{\chi}(z)\,(\mathcal{P}_\cdot - z_0)^{-m}\,(z-\mathcal{P}_\cdot)^{-1}\,dL(z) + \frac{1}{2i\pi}\int_{\Omega_+\cap \mathbf{Im} z=\delta} g(z)\,\widetilde{\chi}(z)\,(\mathcal{P}_\cdot - z_0)^{-m}\,(z-\mathcal{P}_\cdot)^{-1}\,dz.$$

Si $\delta > 0$ est assez petit, sur le domaine d'intégration de la première intégrale, $g = \mathcal{O}(1)$, ce qui permet de montrer que la contribution de ce terme à la trace est $\mathcal{O}\left(h^{-n\sharp}\right)$. La deuxième intégrale est, modulo une erreur du même ordre, égale à

$$J_\cdot = \frac{1}{2i\pi}\int_\gamma g(z)\,(\mathcal{P}_\cdot - z_0)^{-m}\,(z-\mathcal{P}_\cdot)^{-1}\,dz,$$

où γ est un segment $[a,b]$, $\mathbf{Im} a = \mathbf{Im} b = \delta$, $J_+ \ni \mathbf{Re} a < \inf I_+$, $J_+ \ni \mathbf{Re} b > \sup I^+$. On montre alors qu'on peut déformer l'opérateur $\mathcal{P}$ sans changer $J_\cdot$:

$$J_\cdot = \frac{1}{2i\pi}\int_\gamma g(z)\,(\mathcal{P}_{\theta,\cdot} - z_0)^{-m}\,(z-\mathcal{P}_{\theta,\cdot,\cdot})^{-1}\,dz.$$

Soit γ' un contour dans $\Omega\setminus W$ joignant a à b. On montre ensuite (et c'est la partie la plus technique de la preuve) que

$$\mathbf{Tr}\Big[\Big(\frac{1}{2i\pi}\int_{\gamma'} g(z)\,(\mathcal{P}_{\theta,\cdot} - z_0)^{-m}\,(z-\mathcal{P}_{\theta,\cdot})^{-1}\,dz\Big)\Big]_0^1 = \mathcal{O}\left(h^{-n\sharp}\right),$$

ce qui implique

$$\begin{aligned}\mathbf{Tr}\left(\left[J\right]_0^1\right) &= \mathbf{Tr}\left(\Big[\frac{1}{2i\pi}\int_{\gamma\cup\gamma'} g(z)\,(\mathcal{P}_{\theta,\cdot} - z_0)^{-m}\,((z-\mathcal{P}_{\theta,\cdot})^{-1}\,dz\Big]_0^1\right) + \mathcal{O}\left(h^{-n\sharp}\right),\\ &= \mathbf{Tr}\left(\Big[\frac{1}{2i\pi}\int_{\gamma\cup\gamma'} g(z)\,(P - z_0)^{-m}\,(z-\mathcal{P}_\theta)^{-1}\,dz\Big]_0^1\right) + \mathcal{O}\left(h^{-n\sharp}\right),\end{aligned}$$

et le théorème des résidus fournit la contribution des résonances à la formule de trace.

3.2. Application à l'étude des résonances pour des opérateurs de Schrödinger semi-classiques

Soit $\mathcal{P}_1 = -h^2\Delta + V_1(x)$ un opérateur de Schrödinger vérifiant les hypothèses **H1–6**. Pour $E > 0$, on note $\nu_{+,1}(E) = \int_{V_1(x)\geq E} dx$ et on rappelle que le support singulier analytique d'une distribution f, noté $\mathrm{supp}\,\mathrm{sing}_a(f)$, est l'ensemble des points au voisinage desquels cette distribution n'est pas une fonction analytique réelle. On a alors le résultat suivant qui montre l'existence de nombreuses résonances :

THÉORÈME 2 (Sjöstrand 96). — *Soit* $0 < E_0 \in \mathrm{supp}\,\mathrm{sing}_a(\nu_{+,1})$. *Alors pour tout voisinage complexe de* E_0, W, *il existe* $h_0, C > 0$ *tels que pour tout* $0 < h < h_0$ *le nombre de résonances de* $\mathcal{P}$ *dans* W *est supérieur à* Ch^{-n}.

REMARQUE 3.3. — Dans la mesure où la distribution ν_+ est nulle pour

$$|x| > \sup_x |V_1(x)|,$$

son support singulier analytique n'est pas vide.

Pour la démonstration du théorème 2 nous suivons [13] : on commence par construire un deuxième opérateur $\mathcal{P}_0 = -h^2\Delta + V_0$ vérifiant les hypothèses (**H1–6**) et tel que

(1) la mesure $\nu_{+,0}$ est nulle au voisinage de E_0,

(2) l'opérateur $\mathcal{P}_0$ n'a pas de résonance au voisinage de E_0,

(3) $|V_1(x) - V_0(x)| \leq \frac{C}{(1+|x|)^{-n-1}}$.

Cette construction se fait simplement par troncature et régularisation du potentiel.

Pour $\pm E > 0$, on note $\nu_{\pm,\cdot} = \int_{\pm V(x)\geq \pm E} dx$ et $\mu_\pm = \mp \frac{d}{dE}\nu_{\pm,\cdot}$. Les distributions μ_+ et μ_- sont deux mesures positives sur $\mathbb{R}^+$ et $\mathbb{R}^-$ respectivement. Soit μ la distribution d'ordre -1, à support compact définie par

$$\langle \mu, \varphi \rangle = \int (\varphi(V_1(x)) - \varphi(V_0(x)))\, dx, \ \forall \varphi \in C_0^\infty(\mathbb{R}).$$

On a $\mu|_{\mathbb{R}^\pm} = \left[\mu_{\pm,\cdot}\right]_0^1$. Enfin, on note ω la distribution définie par

$$\langle \omega, \varphi \rangle = \iint \left[f\left(\xi^2 + V_\cdot(x)\right)\right]_0^1 dx.d\xi.$$

Des calculs standard de distributions permettent d'exprimer μ et ω en fonction l'une de l'autre et de montrer en particulier que leurs supports singuliers analytiques coïncident. En particulier, $E_0 \in \operatorname{supp} \operatorname{sing}_a(\omega)$. Plus précisément, puisque la distribution ω est réelle, les deux points $(E_0, \pm 1)$ appartiennent au front d'onde analytique de ω. On rappelle la caractérisation du front d'onde analytique, $SS_a(f)$, d'une distribution $f \in \mathcal{D}'(\mathbb{R}^d)$, en termes de transformation de F.B.I. (voir [12]) :

DÉFINITION 3.4. — *Soient $\varrho_0 = (x_0, \xi_0) \in \mathbb{R}^{2d}$ et $\chi \in C_0^\infty(\mathbb{R}^d)$ égale à 1 au voisinage de x_0. On dit que ϱ_0 n'appartient pas au front d'onde analytique de f ($\varrho_0 \notin SS_a(f)$) si et seulement s'il existe V un voisinage dans $\mathbb{C}^d$ de $x_0 - i\xi_0$ et $C, \varepsilon > 0$ tels que*

$$T(\chi f, \lambda, z) = \int e^{-\lambda(z-x)^2/2} \chi f(x)\, dx,$$

vérifie $|T(\chi f, \lambda, z)| \leq C e^{\lambda(\mathrm{Im} z)^2/2 - \varepsilon\lambda}$ sur $[1, +\infty[_\lambda \times V_z$. Cette définition est indépendante du choix de χ.

Puisque $(E_0, \pm 1)$ appartiennent au front d'onde analytique de ω, il existe donc des suites $(\alpha_j, \beta_j)_{j\in\mathbb{N}}$ et $(\varepsilon_j, \lambda_j)_{j\in\mathbb{N}}$ convergeant vers $(E_0, 1)$ et $(0, +\infty)$ respectivement et vérifiant

$$|T(\chi f, \lambda_j, \alpha_j - i\beta_j)| = \left| \int e^{i\lambda_j\left(\beta_j(\alpha_j - E_0) + i(\alpha_j - E_0)^2/2\right)} \chi f(x)\, dx \right| \geq e^{-\varepsilon_j \lambda_j}. \tag{3.3}$$

Soient $a, b > 0$ et

$$\Omega =]E_0 - b, E_0 + b[\times i] - a, a], \; W =]E_0 - b/2, E_0 + b/2[\times i] - a/2, a].$$

On note I et J les intersections respectives de W et Ω avec l'axe réel et $\chi \in C_0^\infty(J)$ égale à 1 sur I. Soient

$$f_j(E) = e^{i\lambda_j\left(\beta_j(\alpha_j - E) + i(\alpha_j - E)^2/2\right)},$$

qui vérifient, si a, b et a/b sont petits, $f_j\,|_{\Omega \setminus W}\,| \leq e^{-\lambda_j/C}$. La formule de trace du théorème 1 donne

$$\begin{aligned} \mathrm{Tr}\Big[(f_j\chi)(\mathcal{P}_\cdot)\Big]_0^1 &= \Big[\textstyle\sum_{\lambda \in \mathrm{Rés}(\mathcal{P}_\cdot) \cap W_-} f(\lambda)\Big]_0^1 + \mathcal{O}(1)\, h^{-n} e^{-\lambda_j/C}, \\ &= \textstyle\sum_{\lambda \in \mathrm{Rés}(\mathcal{P}_1) \cap W_-} f(\lambda) + \mathcal{O}(1)\, h^{-n} e^{-\lambda_j/C}. \end{aligned} \tag{3.4}$$

Pour calculer la trace on peut appliquer une autre formule due à D. Robert [11]

$$\mathrm{Tr}\Big[(f_j\chi)(\mathcal{P}_\cdot)\Big]_0^1 = \frac{1}{(2\pi h)^n} \int (f_j\chi)(E)\, \omega(E)\, dE + \mathcal{O}_j\left(h^{1-n}\right). \tag{3.5}$$

Cette formule se démontre assez simplement en utilisant le calcul symbolique des opérateurs (voir [13], pages 11 et 12 pour un schéma de preuve). D'après (3.4) et (3.5), on obtient

$$\sum_{\lambda \in \mathrm{Rés}(\mathcal{P}_1) \cap W_-} f(\lambda) = \frac{1}{(2\pi h)^n} \int (f_j\chi)(E)\, \omega(E)\, dE + \mathcal{O}(1)\, h^{-n} e^{-\lambda_j/C} + \mathcal{O}_j\left(h^{1-n}\right),$$

donc d'après (3.3)

$$\Big| \sum_{\lambda \in \mathrm{Rés}(\mathcal{P}_1) \cap W_-} f(\lambda) \Big| \geq \frac{e^{-\varepsilon_j \lambda_j} - \mathcal{O}(1)\, e^{-\lambda_j/C}}{(2\pi h)^n} + \mathcal{O}_j\left(h^{1-n}\right).$$

Fixant d'abord j grand puis h petit, on obtient

$$\Big| \sum_{\lambda \in \mathrm{Rés}(\mathcal{P}_1) \cap W_-} f(\lambda) \Big| \geq \frac{1}{Ch^n},$$

ce qui implique le théorème 2 puisque $|f_j| \leq 1$ sur W_- (si on a choisi W assez petit).

4. RÉSONANCES ET QUASI-MODES

L'objet de cette section est de présenter une autre approche permettant de mettre en évidence la présence de résonances. Il s'agit de résultats de toute autre nature que les précédents. Dans la section 3.1, on a montré que les formules de traces permettaient de minorer le nombre de résonances dans un voisinage complexe d'une énergie E_0, c'est-à-dire relativement loin du réel. On aurait pu donner d'autres résultats de J. Sjöstrand et M. Zworski permettant de minorer ce nombre dans des régions de la forme $|\mathbf{Re} z - E_0| < \varepsilon$,

$|\mathbf{Im} z| < Ch^2 |\log h|$. Ici, nous allons mettre en évidence (dans des cas particuliers) la présence de résonances convergeant vers le réel quand h tend vers 0 plus vite que tout polynôme en h^{-1}. Cette dernière propriété est à la fois un avantage et un inconvénient. C'est un inconvénient puisqu'elle restreint la généralité et que la construction que nous présentons nécessite des hypothèses très fortes (permettant de construire des quasi-modes), mais c'est aussi un avantage car les résonances ainsi construites sont par nature très proches du réel et d'un point de vue physique (voire expérimental), ce sont ces résonances qu'il est possible de mesurer. Nous suivons ici la présentation de P. Stefanov [20].

DÉFINITION 4.1. — *Pour toute suite $(h_l)_{l\in\mathbb{N}}$ de réels strictement positifs tendant vers 0 quand l tend vers l'infini, on appelle* quasi-modes (u_j) *associés aux* quasi-résonances (E_j), *toute famille $(u_j(h_l), E_j(h_l))_{j=1}^{m(l)}$ telle qu'il existe une fonction $R(h) = \mathcal{O}\left(h^{-N}\right)$, $\forall N > 0$ et*

- $u_j(h_l) \in \mathcal{H}$, $\forall l, j$,
- $E_j(h_l) \in \mathbb{R}$, $\forall l, j$,
- *Pour tout $l \in \mathbb{N}^*$ et tout $1 \leq j \leq m(l)$,*

$$\| (\mathcal{P} - E_j(h_l)) u_j(h_l) \|_{\mathcal{H}} \leq R(h_l),$$
$$\left| (u_i(h_l), u_j(h_l))_{\mathcal{H}} - \delta_{i,j} \right| \leq R(h_l),$$
$$\exists R_1 > R_0;\ u_j(h_l) = 1_{B(0,R_1)} u_j(h_l).$$

On a alors le résultat suivant

THÉORÈME 3 (Stefanov-Vodev 95, Tang-Zworski 98, Stefanov 98). — *Soient $\mathcal{P}$, un opérateur vérifiant les hypothèses* **H1–6**, *$(h_l) \to 0$, $l \to +\infty$ et $(u_j(h_l), E_j(h_l))$ des quasi-modes. On suppose de plus qu'il existe deux fonctions $0 < a_0 \leq a(h) \leq b(h) \leq b_0 < +\infty$ telles que pour tous $0 \leq j \leq m(l)$, $E_j(h_l) \in [a(h_l), b(h_l)]$. Alors il existe $D > 0, L \in \mathbb{N}$ tel que pour tout $l \geq L$ et toute fonction positive $S(h)$ vérifiant*

$$S(h) \geq \max\{h^{-n^4-1} R(h), De^{-D/h}\},$$

l'opérateur $\mathcal{P}(h_l)$ a au moins $m(l)$ résonances (comptées avec leur multiplicité) dans l'ensemble

$$E = \left[a(h_l) - 6h_l^k, b(h_l) + 6h_l^k\right] \times i\left[-2S(h_j) h_l^{-n^4-1}, 0\right]. \tag{4.1}$$

En d'autres termes, si l est assez grand, on aura au moins autant de résonances proches de $E_j(h_l)$ que de quasi-modes et les résonances correspondantes seront d'autant plus proches du réel que la fonction R est petite, sans pouvoir être plus proches qu'exponentiellement. Au sujet de l'impossibilité pour les résonances de s'approcher du réel plus vite qu'exponentiellement, on pourra consulter [4].

La preuve de ce résultat repose essentiellement sur les deux lemmes suivants dus à S.H. Tang-M. Zworski [22]. Le premier est une estimation *a priori* sur la résolvante en dehors d'un voisinage des résonances :

LEMME 4.2. — *Soient θ vérifiant les conclusions de l'hypothèse* **H6** *et $\chi \in C_0^\infty(\mathbb{R}^d)$ égale à 1 au voisinage de $B(0,R_0)$. Pour tout compact simplement connexe $\widetilde{\Omega}$ vérifiant*

$$\widetilde{\Omega} \subset S_\theta = \{z \in \mathbb{C};\ \max(-\pi, 2\theta - 2\pi) < -\arg z < 2\theta\}$$

et pour toute fonction positive $g(h)$ définie pour $0 < h < h_0$ et assez petite, il existe des constantes $A, h_1 > 0$ telles que pour tout $0 < h < h_1$ et tout

$$z \in \widetilde{\Omega} \setminus \cup_{z \in Rés(\mathcal{P}(h)) \cap \widetilde{\Omega}} \overline{B(z, g(h))},$$

on a

$$\|\chi(\mathcal{P}(h) - z)\chi\|_{\mathcal{L}(\mathcal{H})} \leq A e^{Ah^{-n^\sharp} \log(1/g(h))};$$

et le deuxième lemme est un principe du maximum :

LEMME 4.3. — *Soient (h_l) une suite convergeant vers 0 quand l tend vers l'infini, $H = \cup_l \{h_l\}$ et $F(z,h); h \in H$, une famille de fonctions holomorphes en z définie au voisinage de*

$$\Omega(h) = [E(h) - 5h^k, E(h) + 5h^k] + i\left[-S(h)h^{-n^\sharp - 1}, S(h)\right]. \tag{4.2}$$

On suppose qu'on a

$$|F(z,h)| \leq A e^{Ah^{-n^\sharp} \log(1/hS(h))} \text{ sur } \Omega(h),$$
$$|F(z,h)| \leq 1/|\mathbf{Im} z| \text{ sur } \Omega(h) \cap \{\mathbf{Im} z > 0\}.$$

Alors il existe $h_1 > 0$ et B ne dépendant que de S, A et k tels que pour tout $0 < h < h_1$, $h \in H$,

$$|F(z,h)| \leq B/S(h) \text{ sur } \left[E(h) - h^k, E(h) + h^k\right].$$

On revient à la preuve du théorème 3. On commence par supposer que le nombre de quasi-modes vérifie $m(h) \leq Ch^{-n^\sharp}$. On fixe $h \in \{h_l; l \in \mathbb{N}^*\}$. On note $z_1(h), \ldots, z_{M(h)}(h)$ les pôles de l'opérateur $\chi(\mathcal{P}(h) - z)\chi$ inclus dans l'ensemble E défini par la relation (4.1) et

$$A^{(j)}(h) = \frac{1}{2i\pi} \int_{|z - z_j(h)| = \varepsilon_j} \chi(\mathcal{P}(h) - z)\chi dz,$$

le projecteur spectral correspondant (si ε_j est assez petit). On rappelle que la multiplicité de l'opérateur $A^{(j)}$ est égale à la multiplicité de la résonance z_j. Soit $\Pi(h)$ la projection orthogonale sur l'espace vectoriel engendré par la réunion des images des opérateurs $A^{(j)}$. Alors la dimension de l'image de $\Pi(h)$ est supérieure au nombre de résonances dans l'ensemble E. Nous allons montrer que cette dimension est supérieure au nombre de quasi-modes, $m(h)$. Soit $\Pi'(h) = \mathrm{Id} - \Pi(h)$. On peut montrer par des calculs purement algébriques que l'image de la partie singulière de $\chi(\mathcal{P}(h) - z)\chi$ en un pôle est incluse

dans l'image du projecteur spectral correspondant. L'opérateur $\Pi'(h)\chi(\mathcal{P}(h)-z)\chi$ est donc holomorphe dans E et donc aussi dans l'ensemble

$$\Omega_6(h)=\left[a(h_l)-h_l^k,b(h_l)+h_l^k\right]\times i\left[-2S(h_j)h_l^{-n^\sharp-1},S(h)\right].$$

Soient $E_s(h)=(1-s)a(h)+sb(h)$, $s\in[0,1]$. L'opérateur $\Pi'(h)\chi(\mathcal{P}(h)-z)\chi$ est donc, pour tout $s\in[0,1]$, holomorphe dans l'ensemble

$$\Omega_{5,s}=\left[E_s(h_l)-5h_l^k,E_s(h_l)+5h_l^k\right]\times i\left[-2S(h_j)h_l^{-n^\sharp-1},S(h)\right].$$

On choisit $g(h)=hS(h)$ et d'après le lemme 4.2 on obtient

$$\forall z\in\Omega_6(h)\setminus\cup_{z\in\text{Rés}(\mathcal{P}(h))}\overline{B(z,hS(h))},$$
$$\|\Pi'(h)\chi(\mathcal{P}(h)-z)\chi\|_{\mathcal{L}(\mathcal{H})}\leq Ae^{Ah^{-n^\sharp}\log(1/g(h))}.$$

Les disques dans la relation précédente qui rencontrent $\Omega_{5,s}$ sont inclus dans $\Omega_6(h)$, l'estimation (4.2) est donc vérifiée sur le bord de ces disques donc, par le principe du maximum, sur ces disques et par conséquent sur $\Omega_{5,s}$. Enfin comme $\mathcal{P}(h)$ est autoadjoint, on a pour $\mathbf{Im}z>0$

$$\|\Pi'(h)\chi(\mathcal{P}(h)-z)\chi\|_{\mathcal{L}(\mathcal{H})}\leq\|\chi(\mathcal{P}(h)-z)\chi\|_{\mathcal{L}(\mathcal{H})}\leq\frac{1}{|\mathbf{Im}z|}.$$

et le lemme 4.3 implique donc pour $z\in E_s(h)-h^k,E_s(h)+h^k$, $s\in[0,1]$ (donc pour $z\in[a(h)-h^k,b(h)+h^k]$),

$$\|\Pi'(h)\chi(\mathcal{P}(h)-z)\chi\|_{\mathcal{L}(\mathcal{H})}\leq\frac{B}{S(h)}.$$

Comme pour $z\notin\text{Rés}\mathcal{P}(h)$, $u_j(h)=\chi u_j(h)=\chi(P(h)-z)^{-1}\chi(P(h)-z)u_j$, on a

$$\Pi'(h)u_j(h)=\Pi'(h)\chi u_j(h)=\Pi'(h)\chi(P(h)-z)^{-1}\chi(P(h)-z)u_j. \tag{4.3}$$

Le terme de droite est holomorphe en $z\in E$; par prolongement analytique la relation (4.3) reste donc vraie pour $z=E_j(h)$, ce qui donne

$$\|\Pi'(h)u_j(h)\|_{\mathcal{H}}\leq B\frac{R(h)}{S(h)}\leq Bh^{n^\sharp+1}.$$

Les fonctions $f_j=\Pi(h)u_j(h)$ vérifient donc

$$|(f_j(h),f_i(h))-\delta_{i,j}|\leq Ch^{n^\sharp+1}. \tag{4.4}$$

Comme on a au plus $Ch^{-n^\sharp}$ quasi-modes les fonctions f_j forment d'après (4.4) un système libre et l'image de Π est donc de dimension au moins égale au nombre de quasi-modes. Il reste à supprimer l'hypothèse $m(h)\leq Ch^{-n^\sharp}$. Le raisonnement précédent fonctionnerait encore si on gardait un nombre de quasi-modes de l'ordre $h^{-n^\sharp-\delta}$, $0<\delta<1$, mais alors il montrerait la présence d'un nombre de résonances du même ordre, contredisant l'estimation *a priori* de la remarque 3.2.

RÉFÉRENCES

[1] J. AGUILAR et J. COMBES – *A class of analytic perturbations for one body schrödinger hamiltonians*, Comm. Math. Phys. **22** (1971), 269–279.

[2] V. BABICH et V. BULDYREV – *Short wavelength diffraction theory, asymptotic methods*, Wave phenomena, vol. 4, Springer Verlag, Berlin, 1991, traduit du texte russe de 1972 par E.F. Kuester.

[3] C. BARDOS, J. C. GUILLOT et J. RALSTON – *La relation de Poisson pour l'équation des ondes dans un ouvert non borné. Application à la théorie de la diffusion*, Communications in Partial Differential Equations **7** (1982), 905–958.

[4] N. BURQ – *Décroissance de l'énergie locale de l'équation des ondes pour le problème extérieur et absence de résonance au voisinage du réel*, Acta Mathematica **180** (1998), 1–29.

[5] P. D. LAX et R. S. PHILLIPS – *The time delay operator and a related trace formula*, Topics in functional analysis, Advances in Math. Suppl. Studies **47** (1978), no. 3, 197–395.

[6] V. LAZUTKIN – *Kam theory and semiclassical approximation to eigenfunctions*, Ergebnisse der Mathematik und ihree Grenzgebiete, vol. 24, Springer Verlag, Berlin, 1993, With an addendum by A.I. Shnirelman.

[7] R. MELROSE – *Scattering theory and the trace of the wave group*, Jour. of Func. Anal. **45** (1982), 429–440.

[8] R. MELROSE – *Polynomial bounds on the number of scattering poles*, Jour. of Func. Anal. **53** (1983), 287–303.

[9] G. POPOV – *Quasi-modes for the Laplace operator and glancing hypersurfaces*, Proceedings of conference on microlocal analysis and non linear waves (Minnesota) (J. R. M. Beals, R. Melrose, éd.), Springer Verlag, 1989.

[10] J. RALSTON – *Approximate eigenfunctions for the Laplacian*, Jour. Diff. Geo. **12** (1977), 87–100.

[11] D. ROBERT – *Relative time-delay for perturbation of elliptic operators and semiclassical asymptotics*, Jour. of Func. Anal. **126** (1994), no. 1, 36–82.

[12] J. SJÖSTRAND – *Singularités analytiques microlocales*, Astérisque, vol. 95, Société mathématique de France, 1982.

[13] J. SJÖSTRAND – *A trace formula and application to semi-classical Schrödinger operators*, Séminaire E.D.P. du Centre de Mathématiques de l'École Polytechnique **II** (1996–97).

[14] J. SJÖSTRAND – *A trace formula and review of some estimates for resonances*, Microlocal Analysis and Spectral Theory, NATO ASI series C, vol. 490, Kluwer, 1997, 377–437.

[15] J. SJÖSTRAND – *Resonances for bottles and trace formula*, Prépublication du Centre de Mathématiques de l'École Polytechnique **98-8** (1998).

[16] J. SJÖSTRAND et M. ZWORSKI – *Complex scaling and the distribution of scattering poles*, Journal of the A.M.S. **4** (1991), no. 4, 729–769.

[17] J. SJÖSTRAND et M. ZWORSKI – *Lower bounds on the number of scattering poles*, Comm. in P.D.E. **18** (1993), no. 5-6, 847–857.

[18] J. SJÖSTRAND et M. ZWORSKI – *Lower bounds on the number of scattering poles II*, Jour. of Func. Anal. **123** (1994), no. 2, 336–367.

[19] J. SJÖSTRAND et M. ZWORSKI – *Asymptotic distribution of resonances for convex obstacles*, Prépublication du Centre de Mathématiques de l'École Polytechnique **98-6** (1998).

[20] P. STEFANOV – *Quasimodes and resonances : fine lower bounds*, Preprint (1998).

[21] P. STEFANOV et G. VODEV – *Distribution of resonances for the Neumann problem in linear elasticity outside a strictly convex body*, Duke Math. Jour. **78** (1996), no. 3, 677–714.

[22] S. TANG et M. ZWORSKI – *From quasimodes to resonances*, to appear in Math. Res. Lett. (1998).

Nicolas BURQ

Université Paris-Sud
URA 760 du CNRS
Département de Mathématiques
Bâtiment 425
F-91405 ORSAY Cedex
E-mail : `Nicolas.Burq@math.u-psud.fr`

Astérisque

HENRI CARAYOL

Preuve de la conjecture de Langlands locale pour GL_n : travaux de Harris-Taylor et Henniart

Astérisque, tome 266 (2000), Séminaire Bourbaki,
exp. n° 857, p. 191-243

<http://www.numdam.org/item?id=SB_1998-1999__41__191_0>

© Société mathématique de France, 2000, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (http://smf4.emath.fr/Publications/Asterisque/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Séminaire BOURBAKI Mars 1999
51ème année, 1998-99, n° 857

PREUVE DE LA CONJECTURE DE LANGLANDS LOCALE POUR GL_n : TRAVAUX DE HARRIS–TAYLOR ET HENNIART

par **Henri CARAYOL**

1. INTRODUCTION

Soit K un corps local non archimédien ; il est donc isomorphe soit à un corps p–adique (extension finie de $\mathbf{Q}_p$), soit au corps $\mathbf{F}_q((t))$ des séries de Laurent formelles en une variable sur un corps fini. Nous noterons k le corps résiduel de K, de caractéristique p et de cardinal q. Nous désignerons par $\mathcal{O}$ l'anneau des entiers de K et par ϖ une uniformisante. Enfin, *le groupe de Weil* W_K est le sous-groupe du groupe de Galois $\mathrm{Gal}(\overline{K}/K)$ constitué des éléments qui induisent sur le corps résiduel une puissance entière d'un élément de Frobenius ; notre convention pour les Frobenius dans la suite sera celle dite 'géométrique' : on désignera ainsi un élément de W_K qui induit sur le corps résiduel *l'inverse* d'une substitution de Frobenius $t \to t^q$.

La *théorie du corps de classes local* établit l'existence d'un isomorphisme canonique :

$$W_K^{\mathrm{ab}} \simeq K^*$$

(que nous normaliserons de telle sorte qu'à un Frobenius géométrique corresponde une uniformisante). Cet isomorphisme permet d'identifier l'ensemble des représentations continues irréductibles (bien sûr de degré 1) de $K^* = GL_1(K)$ à l'ensemble des représentations continues de degré 1 de W_K. Comme généralisation non-abélienne de la théorie du corps de classes local, Langlands a conjecturé l'existence pour chaque n de bijections naturelles entre l'ensemble des classes d'équivalence de représentations admissibles irréductibles du groupe $GL_n(K)$ d'une part et l'ensemble des classes d'équivalence de représentations de degré n (Fr–semi-simples) du groupe de Weil–Deligne W_K' d'autre part (voir ci-dessous les définitions plus précises de ces objets). En fait, cette conjecture a été précisée au fil des ans. Les bijections cherchées doivent

être caractérisées par certaines propriétés. On peut faire cela essentiellement de deux façons : soit dans un contexte purement local en termes de *facteurs locaux* L et ϵ, soit en relation avec la *théorie globale* reliant représentations automorphes et représentations de groupes de Galois de corps globaux. Le paragraphe suivant sera consacré à une discussion des propriétés purement locales que l'on attend d'une telle correspondance.

La conjecture locale de Langlands a été l'objet de nombreux travaux depuis le début des années 70. Depuis l'origine, deux voies d'attaque ont été explorées : l'une, purement locale, vise en gros à une description assez explicite des deux types d'objets considérés, en particulier à une classification fine de l'ensemble des représentations irréductibles de GL_n ; une fois les objets classifiés et en faisant éventuellement usage de fonctorialités de Langlands locales, on essaye (et on y parvient dans certains cas) de définir la bijection cherchée et de prouver qu'elle possède toutes les propriétés requises. L'autre voie d'attaque utilise la théorie globale : soit F un *corps global*, c'est-à-dire ou bien un corps de nombres de degré fini sur $\mathbf{Q}$, ou bien un corps de fonctions d'une variable (i.e. de degré de transcendance 1) sur un corps fini. Dans de nombreux cas, on sait associer à des *représentations automorphes* de groupes sur F des représentations (en général ℓ-adiques) du groupe de Galois global $\mathrm{Gal}(\overline{F}/F)$; cela se fait en utilisant les variétés de Shimura (cas des corps de nombres), ou bien (dans le cas des corps de fonctions) les variétés modulaires (de modules elliptiques et de chtoukas) introduites par Drinfeld ainsi que leurs semblables ($\mathcal{D}$–faisceaux elliptiques de Laumon, Rapoport et Stuhler). Si K est l'un des complétés de F, on peut essayer d'utiliser cette construction globale pour en déduire une correspondance locale. Cela marche assez facilement dans le cas des corps de fonctions pour une raison que j'expliquerai dans le paragraphe 3, mais cela ne suffit pas sur un corps de nombres. On peut aussi tenter de procéder dans l'autre sens, partant d'une représentation galoisienne locale, et l'étendant en une représentation galoisienne globale ; dans certains cas, on sait construire une représentation automorphe associée dont on considère alors la composante locale en la place qui nous intéresse ...

La conjecture a d'abord été vérifiée pour GL_2 d'un corps local de caractéristique résiduelle $\neq 2$. Puis, vers le milieu des années 70, Tunnell l'a prouvée pour $GL_2(K)$, avec K le corps $\mathbf{Q}_2$ ou bien une extension contenant les racines cubiques de l'unité, tandis que Deligne l'établissait pour $n = 2$ et K d'égale caractéristique. Tous deux utilisaient des méthodes globales. Le premier résultat général remonte a Kutzko ([Ku]) qui la prouvait pour GL_2 (K quelconque) par des méthodes locales ; ensuite Henniart ([He1]) a démontré le résultat pour GL_3, toujours par des méthodes locales. Il faut

ensuite attendre Laumon, Rapoport et Stuhler ([L-R-S]) en 1991 pour une preuve valide pour tout n, mais limitée aux corps locaux d'égale caractéristique p .

Au début de l'été 1998, Harris et Taylor [H-T] ont prouvé la conjecture pour n quelconque et K un corps p–adique par des méthodes géométriques. Entre-temps différents résultats partiels avaient été obtenus, essentiellement dans des cas 'modérés'. La méthode utilisée par Harris et Taylor est fondée sur une étude fine de la réduction aux mauvaises places des *variétés de Shimura* associées à certains groupes unitaires tordus. Ces variétés permettent d'associer des représentations galoisiennes globales à des formes automorphes, ce qui ne donne pas directement la correspondance voulue. L'essence de leur méthode consiste à établir en gros la chose suivante : si à la représentation automorphe Π correspond la représentation galoisienne Σ, alors la restriction de Σ à la place w que l'on regarde ne dépend (au moins à semi-simplification près) que de la composante locale Π_w de Π et se calcule au moyen d'objets locaux : des représentations définies en termes géométriques des groupes produits $GL_m(K) \times D^*_{m,K} \times W_K$, où $D_{m,K}$ désigne le corps gauche de centre K, de dimension m^2 et d'invariant $1/m$. Ces représentations locales avaient été considérées tout d'abord par Deligne (vers 1970) dans le cas où $m = 2$ et où $K = \mathbf{Q}_p$. Je les avais étudiées ensuite, toujours pour GL_2, dans le cas de corps p–adiques quelconques. Il s'agissait à l'époque d'étudier les courbes modulaires (puis les courbes de Shimura) aux mauvaises places et en particulier de déterminer à ces places les représentations galoisiennes associées. Langlands avait commencé cette étude pour les courbes modulaires, puis Deligne avait achevé le travail en introduisant comme outil fondamental cette représentation locale. J'avais ensuite procédé de façon semblable dans une étude sur les courbes de Shimura. Le travail de Harris et Taylor constitue ainsi une généralisation directe de ces idées à des variétés de Shimura de dimension > 1. De son côté, Drinfeld avait aussi obtenu vers la même époque des objets munis d'actions des groupes produits $GL_m(K) \times D^*_{m,K} \times W_K$, d'où une autre représentation locale. Que ces représentations possèdent la vertu de calculer la correspondance locale de Langlands, en même temps que la correspondance de Jacquet–Langlands locale (reliant représentations de $GL_m(K)$ et de $D^*_{m,K}$) était donc bien plausible et cela était probablement dans l'idée des spécialistes du sujet. J'avais précisé cela sous forme d'une conjecture ([Ca3]). Tout ceci constitue une généralisation non-abélienne de la théorie de Lubin–Tate (que l'on retrouve pour $m = 1$).

Vers Noël 1998, Henniart [He4] a proposé une autre démonstration plus simple et plus courte que celle de Harris et Taylor ; cependant elle va moins loin puisqu'elle

n'établit pas, contrairement à la précédente, la compatibilité aux mauvaises places entre le résultat local et la construction de représentations ℓ–adiques à partir de certaines représentations automorphes. Si l'on veut, le résultat d'Henniart est l'exact analogue de celui de Kutzko pour GL_2 (et du sien dans le cas de GL_3), tandis que celui de Harris et Taylor contient la généralisation des résultats de Langlands, Deligne, et moi-même, dont il a été question ci-dessus. La méthode d'Henniart est aussi de nature globale : elle consiste en gros, à partir d'une représentation du côté galoisien, à la décomposer comme une somme virtuelle d'induites de caractères, puis à considérer dans un groupe convenable la somme correspondante d'hypothétiques 'induites automorphes locales'. Ces dernières, dont on ne sait pas a priori qu'elles existent, proviennent en réalité de constructions globales et le point crucial consiste à prouver que leur formation se localise aussi bien que possible.

Le dénominateur commun entre les deux méthodes se compose d'une part du théorème d'induction de Brauer et d'autre part de l'existence, sous certaines hypothèses, d'une 'induction automorphe non galoisienne'. L'induction automorphe (locale ou globale) est l'opération qui devrait correspondre, d'après la philosophie de Langlands, à l'opération d'induction du côté galoisien. On sait la construire essentiellement pour des extensions cycliques (si l'on excepte certains résultats qui découlent de la théorie des fonctions L), par des méthodes fondées sur la formule des traces d'Arthur–Selberg ; par itération, on les obtient donc aussi pour des extensions *galoisiennes résolubles.* Toutefois Harris avait remarqué dans [Ha2] qu'en combinant l'induction automorphe cyclique et le changement de base avec l'existence des représentations galoisiennes attachées aux formes automorphes, on pouvait attraper certains cas particuliers d'induction automorphe non galoisienne, pour des sous-extensions non galoisiennes dans des extensions galoisiennes résolubles. Cette idée joue un rôle crucial tant dans le travail de Harris et Taylor que dans celui d'Henniart. On remarquera que ceci fait intervenir les variétés de Shimura, puisqu'elles servent à construire les représentations galoisiennes ℓ–adiques. Ces variétés se retrouvent donc aussi, bien que moins visibles, dans le travail d'Henniart. La différence est qu'il ne les utilise qu'aux 'bonnes' places, récupérant ensuite des informations aux mauvaises par des arguments fondés sur les équations fonctionnelles des séries L, au contraire de Harris et Taylor qui attaquent courageusement ces places de mauvaise réduction, suivant en cela le seul principe de la théorie que l'on ait réussi à établir en toute généralité : 'We are in a forest whose trees will not fall with a few timid hatchet blows. We have to take up the double–bitted axe and the cross–cut saw, and hope that our muscles are equal to them' ([La]).

Remarquons enfin que tant le travail de Harris et Taylor que celui d'Henniart ne font qu'établir abstraitement l'existence de la bijection souhaitée. On peut penser que les méthodes locales parviendront sous peu à la rendre plus explicite et aussi à fournir une autre preuve de la conjecture. La théorie correspondante a été considérablement développée ces dernières années, essentiellement par Bushnell, Henniart et Kutzko. En particulier, ils ont construit une bijection explicite entre représentations cuspidales de $GL_n(K)$ et représentations irréductibles du groupe de Weil dans le cas (en principe le plus difficile) où n est une puissance de la caractéristique du corps résiduel, sans pouvoir cependant vérifier que cette bijection satisfaisait toutes les conditions requises. Il m'est impossible de décrire brièvement les derniers développements de cette théorie locale – développements qui justifieraient à eux seuls un autre exposé. C'est pourquoi il n'en sera plus question dans la suite. Je renvoie à [B-H] pour un état actualisé de la question.

Le plan de l'exposé est le suivant : le premier paragraphe est consacré à une description des catégories de représentations locales considérées, ce qui nous permet d'énoncer le théorème principal au paragraphe 2. Le suivant est consacré à un exposé très succinct de quelques points, utiles par la suite, de la théorie de Langlands (surtout globale), en particulier l'idée de l'induction automorphe. Au paragraphe 4 nous expliquons la méthode d'Henniart. Le paragraphe 5 raconte la 'théorie de Lubin–Tate non–abélienne', avec un peu plus d'informations que ce dont on aurait strictement besoin, mais c'est un sujet auquel le conférencier est attaché. Au paragraphe 6, on décrit les variétés de Shimura étudiées par Harris et Taylor, et au suivant on explique le fondement de leur méthode. Finalement, on donne quelques détails dans le dernier paragraphe, concernant surtout la géométrie de la mauvaise réduction. Je n'ai pas eu le temps ni le courage de donner des détails sur l'application des formules des traces de Selberg et de Lefschetz : cette partie du manuscrit repose de façon essentielle sur les résultats de Kottwitz sur les points des variétés de Shimura sur les corps finis, résultats qu'il aurait été trop long d'expliquer ici (et pour lesquels je renvoie à [Cl3]). Cet exposé a d'ailleurs déjà largement dépassé la taille que j'avais initialement prévue. Il a été écrit alors que les travaux en question étaient encore sous forme provisoire et mouvante (en gestation en ce qui concerne celui d'Henniart). Cela a entraîné de nombreux remaniements et revirements, et probablement quelques erreurs, imprécisions ou redites, pour lesquelles je sollicite l'indulgence du lecteur.

1. PRÉSENTATION DES ACTEURS

1.1. Le côté de GL_n

Pour chaque entier $n \geq 1$, le groupe $GL_n(K)$ est localement compact et totalement discontinu. Une représentation π de ce groupe sur un espace vectoriel complexe V est *lisse* si le stabilisateur de chaque vecteur est ouvert. Elle est dite *admissible* si de plus, pour chaque sous-groupe ouvert $H \subset GL_n(K)$, l'espace des invariants V^H est de dimension finie ; cela est d'ailleurs automatique pour π lisse et irréductible. Nous noterons $\mathcal{A}_n(K)$ l'ensemble des classes d'équivalence de représentations admissibles irréductibles de $GL_n(K)$. Remarquer que ces représentations sont en général de dimension infinie (sinon de dimension 1 et factorisées à travers le déterminant). Remarquer aussi qu'il s'agit d'une notion purement algébrique (la topologie de $\mathbf{C}$ ne jouant aucun rôle), de sorte que l'on peut se permettre de remplacer $\mathbf{C}$ par un corps algébriquement clos de caractéristique 0. En fait, dans la suite, nous considérerons souvent des représentations à valeurs dans la clôture algébrique $\overline{\mathbf{Q}}_\ell$ de $\mathbf{Q}_\ell$.

Pour $\pi \in \mathcal{A}_n(K)$, le centre de $GL_n(K)$ (identifié à K^*) agit sur l'espace de π par un (quasi-)caractère que l'on appelle le *caractère central* de π. D'autre part, pour χ un quasi-caractère de K^*, on considère souvent la *tordue* $\pi \otimes \chi$: par définition $\pi \otimes \chi = \pi \otimes (\chi \circ \det)$.

Une façon de construire des représentations de $GL_n(K)$ est l'*induction* : partant d'une partition $n = n_1 + n_2 + \cdots + n_r$ et de représentations π_i des groupes $GL_{n_i}(K)$, on peut voir le groupe produit $GL_{n_1}(K) \times GL_{n_2}(K) \times \cdots \times GL_{n_r}(K)$ comme sous-groupe de Levi du parabolique supérieur $P_{n_1,n_2,\cdots n_r}$ de $GL_n(K)$ associé à la partition choisie de n. Le produit tensoriel des π_i s'étend de manière évidente en une représentation de $P_{n_1,n_2,\cdots n_r}$; par induction unitaire (c'est-à-dire qu'on multiplie avant induction par l'inverse $\delta^{-1}_{P_{n_1,n_2,\cdots n_r}}$ du module), on obtient alors une représentation de $GL_n(K)$ habituellement notée $\pi_1 \times \pi_2 \times \cdots \times \pi_r$. Une représentation irréductible de $GL_n(K)$ est dite *cuspidale* si elle ne figure jamais comme sous-quotient d'une telle induite (à partir d'un sous-groupe parabolique propre).

On montre que toute représentation irréductible π est un sous-quotient de $\pi_1 \times \pi_2 \times \cdots \times \pi_r$ avec π_i des représentations irréductibles cuspidales de $GL_{n_i}(K)$, pour certains entiers n_i de somme égale à n. Ces entiers ainsi que les π_i sont uniquement déterminés à l'ordre près par π. D'autre part, une induite de cuspidales irréductibles $\pi_1 \times \pi_2 \times \cdots \times \pi_r$, à défaut d'être toujours irréductible, est du moins de longueur finie et la suite de ses sous-quotients irréductibles ne dépend pas de l'ordre des π_i. Bernstein et Zelevinsky ont classifié de manière beaucoup plus précise les représentations

admissibles irréductibles de $GL_n(K)$ et décomposé les induites $\pi_1 \times \pi_2 \times \cdots \times \pi_r$: voir par exemple l'exposé de Rodier ([R]) pour plus de détails.

Nous noterons $\mathcal{A}_n^0(K) \subset \mathcal{A}_n(K)$ le sous-ensemble constitué des classes d'équivalence de représentations *cuspidales.*

1.2. Le côté galoisien

On considère, d'un autre côté, le groupe de Weil W_K. Afin de pouvoir décrire les représentations galoisiennes locales ℓ–adiques de façon indépendante de la topologie de $\mathbf{Q}_\ell$, (voir [De2] pour des explications plus précises), on est amené à compliquer un tant soit peu le jeu et à introduire un '*groupe de Weil–Deligne*' W_K' qui possède différentes apparences mais qui de toute façon n'intervient que par la catégorie de ses représentations : une telle représentation (sur $\mathbf{C}$ ou plus généralement sur un corps de caractéristique 0) est un couple (ρ, N) constitué d'une représentation continue ρ de W_K dans un espace de dimension finie E (vu avec sa topologie discrète) et d'un endomorphisme nilpotent N de E qui commute à l'image par ρ du groupe d'inertie, et tel que l'on ait, pour Fr un élément de Frobenius 'géométrique', la relation : $\rho(\mathrm{Fr})N = q^{-1}N\rho(\mathrm{Fr})$. On dit qu'une telle représentation est 'Fr–semi-simple' si ρ est semi-simple.

Nous noterons $\mathcal{G}_n(K)$ l'ensemble des classes d'équivalence de représentations Fr–semi-simples de dimension n complexes (même remarque que ci–dessus) du groupe W_K' ; le sous-ensemble constitué de celles pour lesquelles ρ est *irréductible* (ce qui entraîne $N = 0$) sera noté $\mathcal{G}_n^0(K)$.

1.3. Facteurs locaux

1.3.1. Du côté de GL_n, on définit pour chaque représentation complexe admissible irréductible π de $GL_n(K)$ des facteurs locaux $L(\pi, s)$ et $\epsilon(\pi, s, \psi)$, où $s \in \mathbf{C}$ et où ψ est un caractère complexe additif non trivial fixé de K. Ce sont des facteurs eulériens : $\epsilon(\pi, s, \psi)$ est le produit d'une constante et de $q^{-fs-f(\psi)ns}$, où $f(\psi)$ est le plus grand entier ν tel que ψ soit trivial sur $\mathfrak{p}^{-\nu}\mathcal{O}$ et f un entier que l'on appelle le *conducteur* de π. Quant à $L(\pi, s)$, c'est l'inverse d'un polynôme en q^{-s} de terme constant 1. Dans le cas où $n = 1$, ces facteurs avaient été définis dans la thèse de Tate, puis Jacquet et Langlands les avaient obtenus pour $n = 2$, et enfin Godement et Jacquet [G-J] dans le cas général. Ces derniers procèdent de façon assez semblable à ce que faisait Tate : ils considèrent, pour Φ une fonction localement constante à support compact sur $M_n(K)$ et pour c un 'coefficient matriciel' de π, la fonction Zéta

locale :

$$Z(c,s,\phi)=\int_{GL_n(K)} c(g)\Phi(g)|\det(g)|^s \mathrm{d}^*g$$

(intégrale par rapport à une mesure de Haar sur GL_n). Cette fonction Zéta locale converge pour $\mathrm{Re}(s)$ assez grand, et elle se prolonge en une fonction méromorphe sur $\mathbf{C}$ (en fait une fraction rationnelle en q^{-s}). Le facteur $L(\pi,s)$ est alors caractérisé par le fait qu'il est le plus petit facteur eulérien tel que le quotient :

$$\frac{Z(c,s,\phi)}{L(s,\pi)}$$

soit toujours holomorphe pour chaque Φ et chaque coefficient c de π (en quelque sorte l'inverse du 'dénominateur commun') ; enfin la 'constante locale' $\epsilon(\pi,s,\psi)$ est définie par l'équation fonctionnelle suivante :

$$\frac{Z(c^\vee,1-s+\frac{n-1}{2},\hat{\Phi})}{L(\pi^\vee,1-s)}=\epsilon(\pi,s,\psi)\frac{Z(c,s+\frac{n-1}{2},\Phi)}{L(\pi,s)},$$

avec $c^\vee(g)=c(g^{-1})$, $\pi^\vee$ la représentation contragrédiente de π (dont $c^\vee$ est un coefficient), et enfin $\hat{\Phi}$ la transformée de Fourier de Φ (étant entendu que $M_n(K)$ est identifié à son dual par l'application $(x,y)\to\psi\circ\mathrm{tr}(xy)$ et que l'on prend sur $M_n(K)$ la mesure autoduale).

1.3.2. Lorsque $n\leq 3$, une représentation admissible irréductible π de $GL_n(K)$ est déterminée si l'on connaît les facteurs locaux $L(\pi\otimes\chi,s)$ et $\epsilon(\pi\otimes\chi,s,\psi)$ de chacune des tordues de π par un quasi-caractère, mais cela n'est plus vrai si $n\geq 4$. On a cependant alors un résultat analogue en faisant intervenir les *facteurs locaux de paires* de Jacquet, Piatetski-Shapiro et Shalika ([JPPS1] [JPPS2]). Ces derniers ont associé à chaque couple $\pi\in\mathcal{A}_n(K),\pi'\in\mathcal{A}_{n'}(K)$ des facteurs eulériens $L(\pi\times\pi',s)$ et $\epsilon(\pi\times\pi',s,\psi)$ (que l'on ne définira pas ici). D'après Henniart, si $\pi\in\mathcal{A}_n(K)$ est 'générique', elle est caractérisée par ces facteurs lorsque π' décrit $\mathcal{A}_{n-1}(K)$. Cela est en particulier vrai pour π cuspidale et il suffit alors de faire varier π' dans la réunion des $\mathcal{A}^0_{n'}(K)$, avec $1\leq n'\leq n-1$.

1.3.3. On associe aussi des facteurs locaux $L(\sigma,s)$ et $\epsilon(\sigma,s,\psi)$ à une représentation $\sigma=(\rho,N)$ de degré n du groupe de Weil–Deligne. Le plus simple est de définir le facteur L :

$$L(\sigma,s)=\det(1-q^{-s}.\mathrm{Fr}\mid\sigma_N^{I_K}),$$

avec Fr un Frobenius géométrique et $\sigma_N^{I_K}$ le sous-espace de l'espace de la représentation constitué des vecteurs qui sont à la fois fixés par le groupe d'inertie I_K et tués par N.

Par contre, la définition des facteurs ϵ galoisiens est plus subtile et indirecte : pour $n = 1$, on se ramène par la théorie du corps de classe aux quasi-caractères de K^*, et on décide que le facteur ϵ galoisien est égal au facteur défini par Tate. C'est Langlands qui a ensuite compris comment étendre cela aux représentations de dimension quelconque du groupe de Weil d'un corps local K. Il a prouvé par des méthodes purement locales (puis Deligne a retrouvé cela par des méthodes globales) que l'on pouvait associer à une telle représentation ρ un facteur $\epsilon_K(\rho, s, \psi)$, ces facteurs vérifiant trois propriétés qui les caractérisent :

(i) *Additivité* : pour ρ une représentation de W_K, extension d'une représentation ρ' par une autre ρ'', on a :

$$\epsilon_K(\rho, s, \psi) = \epsilon_K(\rho', s, \psi)\epsilon_K(\rho'', s, \psi).$$

(Noter que cela permet de définir $\epsilon_K(\rho, s, \psi)$ pour ρ une représentation *virtuelle* de W_K.)

(ii) *Induction* : pour M une extension finie de K et ρ une représentation virtuelle de degré 0 de W_M, on a :

$$\epsilon_K(\mathrm{Ind}_M^K(\rho), s, \psi) = \epsilon_M(\rho, s, \psi \circ \mathrm{tr}_{M/K}).$$

(iii) *Degré* 1 : pour ρ une représentation de degré 1 de W_K, le facteur $\epsilon_K(\rho, s, \psi)$ est, via l'isomorphisme de la théorie du corps de classes local, celui défini par Tate.

Finalement, pour $\sigma = (\rho, N)$ une représentation du groupe de Weil–Deligne W_K', on définit :

$$\epsilon_K(\sigma, s, \psi) = \epsilon_K(\rho, s, \psi)\det(-\mathrm{Fr} \mid (\sigma^{I_K}/\sigma_N^{I_K})\ ).$$

2. LE THÉORÈME ET QUELQUES CONSIDÉRATIONS ALENTOUR

On désigne toujours par K un corps local non archimédien et on conserve les notations des paragraphes précédents. La conjecture de Langlands locale pour GL_n peut s'énoncer sous la forme du théorème suivant, prouvé dans [H-T] et [He4] dans

le cas où K est un corps p–adique, tandis que son analogue en égale caractéristique p avait été établi dans [L-R-S] :

2.1. THÉORÈME 1

Il existe une famille de bijections $\sigma_{n,K}$ (que nous noterons pour alléger σ_n , voire σ) :

$$\mathcal{A}_n^0(K) \to \mathcal{G}_n^0(K)$$

$$\pi \quad \to \quad \sigma_n(\pi)$$

vérifiant les propriétés suivantes qui les caractérisent :

(i) *Le déterminant de $\sigma_n(\pi)$ correspond par l'isomorphisme de la théorie du corps de classes local au caractère central de π. En particulier, σ_1 est donnée par cet isomorphisme (on identifiera souvent pour alléger $\sigma_1(\chi)$ à χ).*

(ii) (torsion) *Pour tout quasi-caractère $\chi \in \mathcal{A}_1(K)$, on a :*

$$\sigma_n(\pi \otimes \chi) = \sigma_n(\pi) \otimes \sigma_1(\chi).$$

(iii) *Pour chaque paire de représentations $\pi \in \mathcal{A}_n^0(K)$, $\pi' \in \mathcal{A}_{n'}^0(K)$, on a l'identité de facteurs locaux :*

$$L(\pi \times \pi', s) = L(\sigma_n(\pi) \otimes \sigma_{n'}(\pi'), s) \ ; \ \epsilon(\pi \times \pi', s, \psi) = \epsilon(\sigma_n(\pi) \otimes \sigma_{n'}(\pi'), s, \psi).$$

(iv) (contragrédiente) *Pour $\pi \in \mathcal{A}_n^0(K)$, on a : $\sigma_n(\pi^\vee) = \sigma_n(\pi)^\vee$.*

2.2. On peut établir diverses propriétés supplémentaires de ces bijections, la plus notable, qui apparaîtra plus loin comme conséquence de la théorie de Harris et Taylor, étant la compatibilité avec la construction des représentations ℓ–adiques attachées aux formes automorphes sur des formes de groupes unitaires.

Remarque.— Le théorème ci-dessus est énoncé sous l'hypothèse implicite que le corps des coefficients de nos représentations est le corps $\mathbf{C}$ des nombres complexes. La correspondance n'est pas compatible avec les automorphismes de $\mathbf{C}$ si $q^{\frac{n-1}{2}}$ est irrationnel : cela tient au décalage par $\frac{n-1}{2}$ dans les équations fonctionnelles (1.3.1), qui fait apparaître ce facteur. La géométrie des variétés de Shimura fournit plutôt (sur $\overline{\mathbf{Q}}_\ell$) la correspondance (dite 'de Hecke') $r_n(\pi) = \sigma_n(\pi^\vee) \otimes |\ |^{\frac{1-n}{2}}$, qui est bien invariante et garde donc un sens sur tout corps algébriquement clos de caractéristique 0. Malheureusement, cette correspondance locale se comporte mal vis-à-vis de

l'induction et des sommes directes de représentations galoisiennes (contrairement à celle de Langlands : cf. 2.4 ci-dessous). L'irrationalité de la correspondance de Langlands tient également au fait que le module qui sert à l'induction unitaire comporte des puissances de $q^{1/2}$.

2.3. Quelques réductions

2.3.1. L'*unicité* de la famille de bijections satisfaisant ces propriétés est due à Henniart ([He3] ; voir aussi 1.3.2 ci-dessus).

2.3.2. Il résulte de l'égalité des facteurs ϵ que le conducteur d'Artin de $\sigma_n(\pi)$ est égal au conducteur (cf. 1.3.1) de la représentation π.

L'égalité entre le cardinal de l'ensemble (fini) des classes d'équivalence de représentations cuspidales de GL_n de conducteur fixé modulo torsion par un caractère non ramifié et le cardinal de l'ensemble des classes d'équivalence de représentations irréductibles du groupe de Weil de même conducteur, toujours modulo torsion non ramifiée, constitue ce que l'on appelle la *conjecture de Langlands numérique* ; cela avait été prouvé également par Henniart ([He2]) et constitue l'un des ingrédients clés de la preuve du théorème 1.

2.3.3. Pour prouver le théorème, il 'suffit' en fait de construire des applications :

$$\sigma_{n,K} : \mathcal{A}_n(K)^0 \to \mathcal{G}_n(K),$$

vérifiant (i) à (iv) ; en effet ces propriétés entraînent (comme l'avait remarqué Henniart) que l'image de $\sigma_{n,K}$ est incluse dans $\mathcal{G}_n^0(K)$, puis l'injectivité résulte du fait que les représentations cuspidales de GL_n sont déterminées par les facteurs ϵ de paires ; enfin, la surjectivité découle alors de la conjecture numérique.

2.4. Extension en une bijection entre $\mathcal{A}_n(K)$ et $\mathcal{G}_n(K)$

On sait étendre une famille de bijections $\sigma_{n,K}$ comme ci-dessus en une famille de bijections, que nous noterons encore $\sigma_{n,K}$, entre les ensembles $\mathcal{A}_n(K)$ et $\mathcal{G}_n(K)$, vérifiant les mêmes propriétés (i) à (iv). Cela est expliqué par Rodier [R] et repose sur les travaux de Bernstein et Zelevinsky. À vrai dire c'est l'existence de ces bijections entre les ensembles $\mathcal{A}_n(K)$ et $\mathcal{G}_n(K)$ tout entiers qu'avait conjecturée Langlands ; on s'est aperçu assez vite que le cœur du problème consistait à comparer les représentations cuspidales du côté de GL_n aux représentations irréductibles du côté galoisien.

Donnons une idée vague de la construction : à une représentation irréductible π de GL_n, qui apparaît comme sous-quotient d'une induite de cuspidales

$\pi_1 \times \pi_2 \times \cdots \times \pi_r$ (associée à une partition $n = n_1 + n_2 + \cdots + n_r$), on associe $\sigma_n(\pi) = (\rho, N)$, où ρ est simplement $\sigma_{n_1}(\pi_1) \oplus \sigma_{n_2}(\pi_2) \oplus \cdots \oplus \sigma_{n_r}(\pi_r)$. L'opérateur N dépend du sous-quotient que l'on considère et, pour le définir, on doit utiliser toute la finesse de la classification de Bernstein–Zelevinsky.

Cas particulier : l'induite est irréductible si et seulement si il n'existe pas deux entiers i et j distincts entre 1 et r tels que $n_i = n_j$ et que $\pi_j \simeq \pi_i \otimes |\ |$ (torsion par la valeur absolue normalisée de K). Alors $N = 0$, ce qui est d'ailleurs la seule possibilité pour le ρ alors obtenu.

Ce prolongement jouera un rôle important dans la version de Harris et Taylor de la preuve du théorème 1.

3. UN RAPIDE SURVOL DE LA 'PHILOSOPHIE'

3.1. Théorie globale : quelques mots sur les représentations automorphes

Un corps global F étant donné, on désigne par $\mathbf{A}_F$ son anneau d'adèles, produit restreint de tous les complétés F_v. Si G est un groupe réductif sur F, une *forme automorphe* sur $G(\mathbf{A}_F)$ est une fonction sur ce groupe, invariante à gauche par $G(F)$, et vérifiant les propriétés suivantes : elle doit être U-finie pour U un sous-groupe compact maximal de $G(\mathbf{A}_F)$; elle doit être localement constante en les places non archimédiennes et (dans le cas où F est un corps de nombres) on demande qu'elle soit C^∞ aux places réelles et complexes et annulée par un idéal de codimension finie du centre de l'algèbre enveloppante de la partie archimédienne du groupe. On impose enfin des conditions de croissance qui importeront peu dans la suite de cet exposé dans la mesure où nous considérerons surtout des groupes anisotropes modulo le centre.

Sur l'espace des formes automorphes agissent le groupe $G(\mathbf{A}_F^\infty)$ des adèles finies ainsi que l'algèbre enveloppante de l'algèbre de Lie de la partie archimédienne. Un sous-quotient irréductible de l'espace de formes automorphes (pour le produit de ces deux actions) est appelé une *représentation automorphe* (irréductible) du groupe $G(\mathbf{A}_F)$ – à cet abus de langage près que la partie archimédienne n'agit qu'au travers de l'algèbre enveloppante.

Une telle représentation Π se décompose comme un produit tensoriel 'restreint' $\bigotimes \Pi_v$, v décrivant l'ensemble des places de F. Les contributions locales Π_v sont des représentations admissibles irréductibles des différents groupes $G(F_v)$ (de leurs algèbres enveloppantes aux places archimédiennes).

3.2. La 'philosophie de Langlands' prédit qu'à une représentation automorphe Π comme ci–dessus, dont on suppose la composante archimédienne Π_∞ d'un type raisonnable, on doit pouvoir associer un 'motif' sur un corps de nombres, et donc un système compatible ρ_λ de représentations ℓ–adiques d'un groupe de Galois global. Ce programme – que je laisserai vague afin de ne pas démesurément allonger cet exposé – a été réalisé dans certains cas, tant pour des corps de fonctions que pour des corps de nombres. On utilise des objets géométriques (variétés classifiant les modules elliptiques, les chtoukas ou les $\mathcal{D}$–faisceaux elliptiques pour les corps de fonctions, variétés de Shimura dans le cas des corps de nombres) sur lesquels agissent des opérateurs de Hecke (liés au groupe $G(\mathbf{A}^\infty)$) ; la construction consiste à décomposer la cohomologie ℓ–adique des objets considérés sous l'action des opérateurs de Hecke. On prouve ensuite une relation entre Π et ρ_λ du type suivant : pour presque toute place (= toutes sauf un nombre fini) des fonctions L automorphes locales (analogues à celles définies plus haut) coïncident (à un décalage près) avec leur contreparties galoisiennes.

Il est important de signaler ici une différence essentielle entre corps de fonctions et corps de nombres : dans le cas des corps de fonctions, on sait par la théorie de Grothendieck–Deligne–Laumon des fonctions L qu'elles admettent des équations fonctionnelles globales dont la constante est le produit des constantes locales. On sait ceci aussi du côté automorphe pour certains groupes (en particulier les formes de groupes linéaires) et alors des arguments maintenant assez standards (dont nous verrons un exemple ci–dessous en exposant la méthode d'Henniart) permettent souvent d'en déduire l'égalité des facteurs L et ϵ à *chaque* place. Dans les cas des corps de fonctions, l'étude de ce que l'on nomme habituellement 'les mauvaises places', qui correspondent en fait aux places de mauvaise réduction de la variété modulaire, est donc plus ou moins automatique. Cela permet de définir une correspondance locale chaque fois que l'on dispose d'une correspondance globale 'faible' (i.e avec des renseignements portant sur le complémentaire d'un nombre fini de places). Dans le cas des corps de nombres au contraire, on ne connaît à l'heure actuelle aucune façon de prouver a priori que les fonctions L associées à des représentations ℓ–adiques admettent des équations fonctionnelles ; on en est réduit à seulement le constater quand on arrive, après bien des efforts à chacune des mauvaises places, à identifier finalement notre fonction L galoisienne à une fonction L automorphe.

Rendons ces idées un peu plus explicites dans le cas qui nous servira par la suite.

3.3. Résultats de Kottwitz et Clozel (complétés par Taylor)

Soit F un corps CM, que l'on suppose obtenu comme le composé d'un corps totalement réel F^+ et d'un corps quadratique imaginaire. Soit Π une représentation automorphe parabolique (c'est-à-dire en gros qui ne provient pas d'un sous-groupe parabolique propre) du groupe $GL_n(\mathbf{A}_F)$. On suppose que Π est isomorphe à $\overline{\Pi}^\vee$ (où $\overline{\Pi}$ désigne la représentation déduite de Π par composition avec l'automorphisme non trivial de $\mathrm{Gal}(F/F^+)$; on fait aussi l'hypothèse que la composante archimédienne Π_∞ a un caractère infinitésimal régulier ; on suppose enfin qu'il existe une place v de F^+ décomposée dans F telle que Π_v appartienne à la série discrète. Alors on peut associer à Π un système compatible $\rho_{\lambda,\Pi}$ de représentations ℓ–adiques de dimension n du groupe $\mathrm{Gal}(\overline{F}/F)$, caractérisé par la propriété suivante : en presque chaque place finie v de F, la composante Π_v et la restriction de $\rho_{\lambda,\Pi}$ au groupe de Galois local se correspondent au sens de la correspondance de Langlands (triviale pour les places non ramifiées) tordue par $|\ |^{\frac{1-n}{2}}$. Cela est essentiellement dû à Clozel [Cl1], qui, s'appuyant sur des travaux de Kottwitz ([Ko1] [Ko2]), obtenait une représentation de dimension an, puis a été complété par un argument de Taylor (non publié mais expliqué dans [Ha1]) prouvant que la représentation de Clozel était bien (après semi-simplification) une somme de a copies de la représentation voulue. La construction de la représentation de Clozel consiste à 'descendre' Π en une forme automorphe sur une forme de groupe unitaire (analogue à celle que nous rencontrerons plus loin) puis à travailler sur la variété de Shimura associée à ce dernier. Nous étudierons la géométrie de cette construction (y compris aux mauvaises places) très en détail plus loin dans cet exposé, puisqu'elle joue un rôle central dans le travail de Harris et Taylor.

3.4. Fonctorialités (cas cyclique)

Si l'on admet que la conjecture de Langlands locale est vraie, on voit que, pour K'/K une extension de corps locaux, il doit exister une application de 'changement de base' $\mathcal{A}_n(K) \to \mathcal{A}_n(K')$ qui reflète du côté de GL_n l'opération restriction du groupe de Weil–Deligne de K à celui de K' ; d'autre part l'opération d'induction associe à une représentation de degré n de $W'_{K'}$ une représentation de degré nd de W'_K (avec d le degré de l'extension) et cela nous permet de prédire l'existence d'une application ('induction automorphe') de $\mathcal{A}_n(K')$ vers $\mathcal{A}_{nd}(K)$. De plus ces opérations doivent se globaliser en des opérations analogues sur les représentations automorphes.

Le changement de base et l'induction automorphe pour GL_n ont été construites par voie globale (par des méthodes fondées sur la formule des traces de Selberg)

dans le cas d'extensions K'/K *cycliques* (d'où aussi le cas galoisien résoluble) : le changement de base par Arthur et Clozel [A-C] en suivant la voie que Langlands avait explorée pour GL_2 ; l'induction automorphe par Henniart et Herb [H-H], à la suite de travaux de Kazhdan et aussi de résultats plus faibles dus à Arthur et Clozel. Ces applications sont caractérisées par des identités de caractères (tordus) et sont définies indépendamment de la conjecture de Langlands ; on vérifie simplement que ce sont bien les applications voulues là où cette conjecture est triviale, c'est-à-dire aux places non ramifiées.

3.5. Induction automorphe dans certains cas non galoisiens

La théorie dans son état actuel s'avère incapable d'aborder systématiquement l'induction automorphe pour des extensions qui ne sont pas composées d'extensions cycliques. Toutefois Harris a remarqué dans [Ha2] qu'en combinant l'induction automorphe cyclique avec l'existence des représentations galoisiennes décrites ci-dessus (3.3), on pouvait en attraper certains cas particuliers, pour des sous-extensions non galoisiennes dans des extensions galoisiennes résolubles. Cette idée joue un rôle crucial tant dans le travail de Harris et Taylor que dans celui d'Henniart (c'est essentiellement l'intersection des deux méthodes, si l'on excepte l'usage du théorème de Brauer). Expliquons l'idée dans le cas particulier le plus simple d'une extension cubique non galoisienne.

Soit donc F'/F une extension cubique non normale de corps de nombres CM, et soit $\tilde{F}$ la clôture normale de F ; c'est donc un corps de degré 6 sur F. On note F^0 l'unique extension de degré 2 de F contenue dans $\tilde{F}$. On suppose que tous ces corps sont obtenus comme composés de corps totalement réels notés respectivement $F^+, F'^+, F^{0+}, \tilde{F}^+$ et d'un même corps quadratique imaginaire. Donnons-nous un caractère algébrique χ du groupe des classes d'idèles de F' ; on désire construire une représentation automorphe Π de $GL_3(\mathbf{A}_F)$, que l'on puisse considérer comme l'induite automorphe de χ.

Considérons la restriction au groupe de Weil global W_{F^0} de l'induite $\rho = \mathrm{Ind}_{F'}^{F}(\chi)$ du caractère (encore noté χ) correspondant à χ par la théorie du corps de classes global : on voit que cette restriction coïncide avec l'induite ρ^0 de $W_{\tilde{F}}$ à W_{F^0} de la restriction $\tilde{\chi}$ de χ (vu comme caractère de $\mathbf{A}_{\tilde{F}}$, $\tilde{\chi} = \chi \circ N_{\tilde{F}/F'}$). Du côté automorphe, cela nous dit que le changement de base de F à F^0 de notre hypothétique Π doit coïncider avec l'induite automorphe $\Pi^0 = \mathrm{IA}_{\tilde{F}}^{F^0}(\tilde{\chi})$ (que l'on sait exister puisque l'extension $\tilde{F}/F^0$ est cyclique). Il nous reste à descendre cette dernière représentation de F^0 à F. La théorie du changement de base pour l'extension F^0/F nous dit que

c'est possible parce que la représentation considérée est stable par la conjugaison de F^0 relativement à F. Malheureusement, on a deux possibilités Π' et Π'', qui diffèrent par torsion par le caractère d'ordre 2 de $\mathbf{A}_F^*$ défini par l'extension F^0, entre lesquelles on ne sait a priori pas choisir.

Faisons maintenant l'hypothèse supplémentaire que le composé $\overline{\chi}$ de χ avec la conjugaison de F' relativement à F'^+ coïncide avec χ^{-1}. Alors il en est de même pour $\tilde{\chi}$, et Π^0 vérifie la propriété d'autodualité tordue de 3.3 ci-dessus. Supposons de plus qu'elle vérifie aussi les autres propriétés de 3.3, portant sur le caractère infinitésimal et sur l'existence d'une place finie v^0 où la représentation locale appartient à la série discrète. Dans ce cas les résultats de Clozel associent à Π^0 une représentation galoisienne qui n'est autre, après une torsion convenable, que ρ^0! Ajoutons encore parmi nos hypothèses que ρ^0 est irréductible et que v^0 est l'unique place de F^0 divisant une place v de F. Alors on peut montrer que Π' et Π'' vérifient encore les hypothèses de 3.3. Il leur correspond des représentations ρ' et ρ'' (avec la torsion convenable) du groupe $W_{F'}$ de mêmes restrictions ρ^0. Mais une représentation de restriction ρ^0 ne peut être que ρ ou sa tordue par le caractère d'ordre 2 défini par l'extension F^0/F. L'une des deux représentations ρ' ou ρ'', disons ρ', coïncide donc avec ρ : nous pouvons ainsi choisir entre Π' et Π'' et poser $\mathrm{IA}_{F'}^F(\chi) = \Pi'$: c'est l'induite automorphe cherchée. Elle correspond en presque chaque place à l'induite ρ, mais ceci ne nous dit rien de ce qui se passe aux mauvaises places (c'est ce qu'Harris nomme une induite automorphe 'faible'). En particulier on ne sait pas a priori que la composante de $\mathrm{IA}_{F'}^F(\chi)$ en une telle place $\mathfrak{p}$ ne dépend que des composantes $\chi_\mathfrak{q}$ en les places $\mathfrak{q}$ divisant $\mathfrak{p}$. On le saura une fois établis les résultats de Harris et Taylor.

Des arguments semblables permettent souvent, sous des hypothèses techniques compliquées, de construire l'induite automorphe $\mathrm{IA}_{F'}^F(\chi)$ pour des extensions non normales F'/F plus générales (mais contenues dans des extensions résolubles) : voir [Ha2].

3.7. La correspondance de Jacquet–Langlands

Il s'agit d'une des nombreuses fonctorialités prédites par la théorie. On considère l'ensemble des représentations de $GL_n(K)$ (lisses irréductibles) qui sont *essentiellement de carré intégrable*, c'est-à-dire dont les coefficients matriciels sont de carré intégrable modulo le centre. À torsion près, une telle représentation contribue discrètement à la formule de Plancherel du groupe. D'après Bernstein et Zelevinsky, on l'obtient comme l'unique quotient irréductible d'une induite $r \times r|\ | \times r|\ |^2 \times \cdots \times r|\ |^{t-1}$ associée à une représentation cuspidale r de $GL_s(K)$ pour s un diviseur de n. No-

tons $\mathcal{A}_n^d(K)$ leur ensemble ; il contient l'ensemble $\mathcal{A}_n^0(K)$ des cuspidales et n'est pas beaucoup plus gros si n est premier : on ajoute seulement les tordues d'une même représentation, la représentation spéciale. Il y en a un peu plus si n est divisible.

D'autre part, soit D un corps gauche de centre K et de dimension n^2 sur K. Les représentations lisses irréductibles du groupe multiplicatif D^* sont de dimension finie. On note $\mathcal{A}_D$ leur ensemble.

Tant les éléments de $\mathcal{A}_D$ que ceux de $\mathcal{A}_n^d(K)$ admettent des caractères : c'est clair pour les premiers, moins pour les seconds car il s'agit de représentations de dimension en général infinie. On procéde de façon habituelle, définissant d'abord un caractère distribution, dont on montre que sa restriction aux éléments semi-simples réguliers est une fonction.

D'autre part on a une bijection entre les classes de conjugaison régulières de D^* et les classes de conjugaison semi-simples régulières elliptiques de $GL_n(K)$, qui associe entre eux des éléments ayant le même polynôme minimal.

THÉORÈME.— *Il existe une bijection* $\pi \to \mathrm{JL}(\pi)$ *de* $\mathcal{A}_n^d(K)$ *sur* $\mathcal{A}_D$, *caractérisée par l'égalité suivante de caractères, pour* $g \in GL_n(K)$ *semi-simple elliptique régulier et* $d \in D^*$ *possédant le même polynôme minimal* :

$$\Theta_\pi(g) = (-1)^{n-1}\, \Theta_{\mathrm{JL}(\pi)}(d)\ .$$

La démonstration, due à Rogavski, Kazhdan, Vignéras ([Ro], [D-K-V] ; voir aussi [Ba] pour le cas d'égale caractéristique p), procède par voie globale et repose sur la formule des traces de Selberg. Le théorème admet d'ailleurs des variantes globales (correspondance entre représentations automorphes sur différentes formes intérieures de GL_n.)

Par souci d'économie, nous noterons également JL la bijection dans l'autre sens.

4. LA MÉTHODE D'HENNIART

4.1. Sa construction donne l'application $\sigma \to \pi(\sigma)$ des représentations galoisiennes vers les représentations de GL_n, alors que nous verrons dans la suite que Harris et Taylor procèdent plutôt dans l'autre direction.

Notons M_K le $\mathbf{Z}$–module libre de base la réunion $\bigsqcup_{m\geq 1} \mathcal{A}_m^0(K)$. À chaque représentation admissible irréductible π de $GL_n(K)$ on peut associer un élément de

M_K noté $[\pi]$, que l'on appelle son *support cuspidal* et qui n'est rien d'autre que la somme des classes des représentations cuspidales π_i (comptées avec leurs éventuelles multiplicités) telles que π figure dans leur induite $\pi_1 \times \pi_2 \times \cdots \times \pi_r$. On munit ce module du produit scalaire $\langle\ ,\ \rangle$ pour lequel la base qui a servi à le définir est orthonormée.

Réduite à sa plus simple expression, la méthode d'Henniart consiste en ceci : partons de $\sigma \in \mathcal{G}_n^0(K)$; on se ramène aisément à supposer que son déterminant est d'ordre fini, de sorte qu'elle se factorise à travers un groupe de Galois fini. Le théorème d'induction de Brauer permet d'exprimer σ comme une combinaison virtuelle d'induites :

$$\sigma = \sum n_i \mathrm{Ind}_{L_i}^K \chi_i\ ,$$

où les n_i sont des entiers relatifs, les L_i des extensions de K, avec des caractères d'ordres finis χ_i de leurs groupes multiplicatifs (identifiés par la théorie du corps de classes local à des caractères des groupes de Weil correspondants) ; enfin la notation $\mathrm{Ind}_{L_i}^K$ est l'abréviation usuelle pour désigner l'induction de W_{L_i} à W_K.

Les extensions L_i/K ne sont en général pas galoisiennes. Puisqu'il s'agit de corps locaux, on sait du moins que leurs clôtures galoisiennes sont des extensions *résolubles*. Il s'agit alors d'utiliser l'induction automorphe non galoisienne dont j'ai tenté de donner l'idée en 3.5 et d'associer à chaque couple (L_i, χ_i) une représentation 'induite automorphe locale' de $GL_{d_i}(K)$ (avec d_i le degré de l'extension L_i/K) ; la combinaison linéaire dans le groupe M_K des supports cuspidaux de ces 'induites', pondérés des coefficients n_i, doit être la classe de la représentation cherchée $\pi(\sigma)$.

4.2. Cette idée simple se heurte à des obstacles évidents, qu'Henniart parvient toutefois à surmonter. Le plus inquiétant est sans doute le fait que l'induction automorphe locale non galoisienne n'existe pas a priori. En effet l'induction automorphe non galoisienne est définie comme un procédé global ; moyennant certaines précautions, on peut voir χ_i comme la composante locale en une place w_i d'un corps global $\mathbf{L}_i$ d'un caractère $\boldsymbol{\chi}_i$ du groupe des classes d'idèles de $\mathbf{L}_i$; ce corps global doit être choisi de telle sorte qu'il contienne un sous-corps $\mathbf{K}$ d'indice d_i dont le complété en w_i s'identifie à K. On s'arrange d'autre part pour que soient satisfaites les propriétés qui autorisent l'induction automorphe (cf. 3.5) : en particulier, l'extension globale $\mathbf{L}_i/\mathbf{K}$ doit être sous-extension d'une extension galoisienne résoluble et obtenue à partir d'une extension de corps totalement réels par composition avec un corps quadratique imaginaire.

Quant au caractère $\boldsymbol{\chi}_i$, il doit être autodual (son inverse doit coïncider avec son composé par la conjugaison complexe), et vérifier des conditions de régularité à l'infini et à certaines places finies. Sous ces conditions, on peut construire l'induite automorphe $\mathrm{IA}^{\mathbf{K}}_{\mathbf{L}_i}(\boldsymbol{\chi}_i)$. C'est une représentation automorphe irréductible du groupe $GL_{d_i}(\mathbf{A}_{\mathbf{K}})$; sa composante locale $\mathrm{IA}^{\mathbf{K}}_{\mathbf{L}_i}(\boldsymbol{\chi}_i)_{w_i}$ en w_i est une représentation irréductible de $GL_{d_i}(K)$. Moralement, ce doit être l'induite automorphe locale de χ_i, mais on ne sait rien de tel a priori : elle pourrait bien ne pas dépendre seulement de la composante locale χ_i, mais aussi de toutes les données globales que l'on a fixées.

4.3. On remédie partiellement à cela en commençant par ne considérer que des représentations galoisiennes σ factorisées à travers le groupe de Galois d'une extension fixée E/K. Dans ces conditions, Henniart prouve que l'on peut trouver un corps global $\mathbf{K}$ fixe (dépendant de E), admettant K comme complété en une place w et telles que les constructions précédentes puissent être effectuées, si on se limite aux représentations factorisées par $\mathrm{Gal}(E/K)$, en le prenant comme corps de base.

On prolonge donc comme ci-dessus les χ_i en des caractères $\boldsymbol{\chi}_i$ du groupe des classes d'idèles de certaines extensions $\mathbf{L}_i$ de $\mathbf{K}$, qui par la théorie du corps de classes global correspondent à des caractères (complexes continus) du groupe de Weil global de ces extensions ; on peut alors induire ces caractères en des représentations, notées R_i, du groupe de Weil global de $\mathbf{K}$. Par définition, l'induite automorphe $\Pi_i = \mathrm{IA}^{\mathbf{K}}_{\mathbf{L}_i}(\boldsymbol{\chi}_i)$ correspond à R_i, en presque chaque place de $\mathbf{K}$, par la correspondance de Langlands locale (triviale aux places non ramifiées). En particulier, les facteurs locaux L et ϵ du côté automorphe sont égaux aux facteurs galoisiens correspondants, sauf peut-être en un nombre fini de places. Mais alors l'existence d'équations fonctionnelles de part et d'autre, dont la constante globale est le produit des constantes locales, force alors les facteurs de part et d'autre à coïncider en *chaque* place. Nous allons mettre cela en œuvre plutôt pour les facteurs locaux de paires.

Fixons pour cela une autre représentation σ' de $GL_{n'}(K)$ et faisons-lui subir les mêmes constructions qu'à σ, ce qui fait apparaître des extensions locales L'_j et globales $\mathbf{L}'_\mathbf{j}$, des entiers n'_j, des caractères χ'_j et $\boldsymbol{\chi}'_j$, des induites galoisiennes R'_j et automorphes Π'_j. Le fait que ces objets se correspondent comme précédemment en presque chaque place v de $\mathbf{K}$ se traduit en ces places par des égalités de facteurs L et ϵ de paires :

$$L((\Pi_i)_v \times (\Pi'_j)_v \ , \ s) = L((R_i)_v \otimes (R'_j)_v \ , \ s) \ ,$$

$$\epsilon((\Pi_i)_v \times (\Pi'_j)_v \ , \ s \ , \ \psi_v) = \epsilon((R_i)_v \otimes (R'_j)_v \ , \ s \ , \ \psi_v) \ ,$$

où $\psi = \otimes\psi_v$ désigne un caractère du groupe des classes d'idèles de $\mathbf{K}$, dont la composante ψ_w s'identifie à ψ. Les mêmes égalités valent d'ailleurs après passage aux contragrédientes, ou après torsion par un caractère du groupe des classes d'idèles de $\mathbf{K}$. Le point crucial est le fait que toutes les constructions ont été effectuées avec le même corps global de base $\mathbf{K}$.

4.4. Mais la fonction L globale de la paire $\Pi_i \times \Pi'_j$ et celle de la représentation $R_i \otimes R'_j$ (c'est-à-dire le produit des facteurs L locaux correspondants) admettent toutes deux des équations fonctionnelles, dont les constantes globales sont les produits des constantes locales associées. Pour la seconde, cela découle du fait qu'il s'agit de la fonction L associée à une représentation *complexe continue* du groupe de Weil. Par une technique bien connue (développée essentiellement par Henniart) de torsion par des caractères très ramifiés, ce qui fait dégénérer les facteurs ϵ en toute place sauf en celle où on désire les regarder, on obtient alors l'égalité en toute place des 'facteurs γ' locaux, et en particulier à la place w cela donne :

$$\gamma((R_i)_w \otimes (R'_j)_w \ , \ s \ , \ \psi_w) = \gamma((\Pi_i)_w \times (\Pi'_j)_w \ , \ s \ , \ \psi_w) \ ,$$

avec par définition :

$$\gamma((R_i)_w \otimes (R'_j)_w \ , \ s \ , \ \psi_w) = \frac{\epsilon((R_i)_w \otimes (R'_j)_w \ , \ s \ , \ \psi_w) L((R_i)_w^\vee \otimes (R'_j)_w^\vee \ , \ 1-s)}{L((R_i)_w \otimes (R'_j)_w \ , \ s)}$$

et l'expression analogue pour $\gamma((\Pi_i)_w \times (\Pi'_j)_w \ , \ s \ , \ \psi_w)$.

De cette égalité entre facteurs γ, on déduit les égalités entre leurs constituants respectifs :

$$\epsilon((\Pi_i)_w \times (\Pi'_j)_w \ , \ s \ , \ \psi_w) = \epsilon((R_i)_w \otimes (R'_j)_w \ , \ s \ , \ \psi_w) \ ,$$

$$L((\Pi_i)_w \times (\Pi'_j)_w \ , \ s) = L((R_i)_w \otimes (R'_j)_w \ , \ s) \ ,$$

$$L((\Pi_i)_w^\vee \times (\Pi'_j)_w^\vee \ , \ s) = L((R_i)_w^\vee \otimes (R'_j)_w^\vee \ , \ s) \ .$$

Pour pouvoir déduire ces trois égalités de l'égalité des facteurs γ, on utilise le fait que nos représentations $(\Pi_i)_w$ et $(\Pi'_j)_w$ sont 'tempérées'. On a même le résultat plus précis suivant, qui découle de la construction de l'induite automorphe :

Lemme : $(\Pi_i)_w$ *est une induite irréductible* $\pi_1 \times \cdots \times \pi_r$ *de représentations cuspidales unitaires.*

Ceci correspond du côté galoisien au fait que $\mathrm{Ind}_{E_i}^{K}\chi_i$ est une représentation galoisienne semi-simple (car factorisée à travers un groupe fini), somme de représentations irréductibles unitaires. Il résulte du lemme que les expressions des facteurs γ sont des fractions rationnelles en q^{-s} écrites sous forme irréductible.

4.5. Les égalités qui précèdent vont nous servir de façon essentielle à prouver que nos constructions sont indépendantes des choix effectués, à l'exception peut–être du choix de $\mathbf{K}$: il s'agit du choix d'une expression de σ comme combinaison d'induites, choix des corps globaux $\mathbf{E_i}$ et des caractères $\boldsymbol{\chi}_i$.

Revenons à notre idée de départ en considérant l'élément suivant de M_K :

$$\pi(\sigma,\cdots)=\sum n_i[(\Pi_i)_w]\ .$$

Cet élément dépend des nombreux choix que nous avons effectués (et qui sont rappelés par les points de suspension dans la formule ci-dessus). La construction analogue en partant de σ' nous fournit

$$\pi(\sigma',\cdots)=\sum n'_j[(\Pi'_j)_w]\ .$$

De l'égalité des facteurs L locaux va résulter le résultat crucial suivant :

PROPOSITION 4.5.— *Le produit scalaire* $\langle\pi(\sigma,\cdots)\ ,\ \pi(\sigma',\cdots)\rangle$ *est égal au nombre d'entrelacements* $\dim.(Hom_{W_K}(\sigma,\sigma'))$.

Preuve : On sait que les fonctions L locales correspondant à une représentation galoisienne irréductible sont holomorphes en $s=0$, à l'exception près de la représentation triviale pour laquelle on a un pôle simple. La dimension de l'espace des entrelacements entre σ et σ' est donc égal à l'ordre du pôle en $s=0$ de la fonction L locale :

$$L(\sigma\otimes\sigma'^{\vee},s)\ ==\ \prod_{i,j}L\big((R_i)_w\otimes(R'_j)^{\vee}_w,s\big)^{n_in'_j}.$$

Souvenons-nous d'autre part que, d'après le lemme ci–dessus, $(\Pi_i)_w$ (resp. $(\Pi'_j)_w$) est une induite de cuspidales unitaires irréductibles $\pi_{i,1}\times\cdots\times\pi_{i,r}$ (resp. $\pi'_{j,1}\times\cdots\times\pi'_{j',r'}$), dont la somme dans le groupe M_K constitue le support $[(\Pi_i)_w]$ (resp. $[(\Pi'_j)_w]$). Les facteurs locaux de telles induites se calculent comme produits des facteurs des constituants, et il en est de même pour les facteurs locaux de paires. L'ordre du pôle en 0 de $L\big((R_i)_w\otimes(R'_j)_w{}^{\vee},s\big)=L\big((\Pi_i)_w\times(\Pi'_j)_w{}^{\vee},s\big)$ est donc celui du produit :

$$\prod_{\alpha,\beta}L\big(\pi_{i,\alpha}\times\pi'_{j,\beta}{}^{\vee},s\big)^{n_in'_j}\ .$$

Enfin il est connu que, pour un tel couple $\pi_{i,\alpha} \times {\pi'_{j,\beta}}^{\vee}$ de cuspidales irréductibles, l'ordre du pôle est ce qu'on attend par analogie avec le côté galoisien, c'est-à-dire 1 si $\pi_{i,\alpha}$ est isomorphe à $\pi'_{j,\beta}$ et 0 sinon. L'ordre du pôle du produit ci-dessus est donc égal au produit de $n_i n'_j$ et du produit scalaire $\langle\, [(\Pi_i)_w]\ ,\ [(\Pi'_j)_w]\,\rangle$.

Il en résulte finalement que la fonction $L(\sigma \otimes \sigma'^{\vee}, s)$ admet un pôle d'ordre

$$\sum n_i n'_j \langle\, [(\Pi_i)_w]\ ,\ [(\Pi'_j)_w]\,\rangle \ =\ \langle \pi(\sigma, \cdots)\ ,\ \pi(\sigma', \cdots)\rangle\ ,$$

ce qu'il fallait démontrer.

4.6. La proposition ci–dessus admet la conséquence suivante, qui nous rapproche considérablement du but :

PROPOSITION 4.6.— (i) *Pour $\sigma \in \mathcal{G}_n^0(K)$ (toujours supposée factorisée via $Gal(E/K)$), la représentation $\pi(\sigma, \cdots)$ ne dépend pas à isomorphisme près des choix effectués concernant la décomposition de σ en combinaison linéaire d'induites, les extensions globales $\mathbf{L}_i$ et le prolongement des χ_i en caractères idéliques ; par contre elle peut a priori dépendre du choix de $\mathbf{K}$. Nous la noterons désormais $\pi(\sigma, \mathbf{K})$.*

(ii) *Sous la même hypothèse, $\pi(\sigma, \mathbf{K})$ est la classe dans M_K d'un élément bien déterminé de $\mathcal{A}_n^0(K)$ que nous noterons encore $\pi(\sigma, \mathbf{K})$.*

(iii) *L'application qui à σ (factorisée par $Gal(E/K)$) associe $\pi(\sigma, \mathbf{K})$ est compatible au passage aux contragrédientes, à la torsion par un caractère de $Gal(E/K)$. Elle fait se correspondre les caractères centraux du côté automorphe et les déterminants du côté galoisien. Elle préserve les facteurs locaux de paires.*

Preuve : On commence par prouver le second point : pour σ irréductible, on applique la proposition précédente avec $\sigma = \sigma'$ et en faisant la construction pour un choix fixé des données $n_i, \chi_i, \mathbf{L}_i, \boldsymbol{\chi}_i$. On obtient $\pi \in M_K$ qui vérifie :

$$\langle \pi, \pi \rangle \ =\ \dim.(Hom(\sigma, \sigma)) \ =\ 1\ ,$$

ce qui force π à être à un signe ± 1 près la classe d'un élément de $\mathcal{A}_m^0(K)$ pour un certain entier m. On vérifie que le signe est $+1$ et que $m = n$ en utilisant l'application linéaire λ de M_K dans $\mathbf{Z}$ qui envoie tout élément de $\mathcal{A}_d^0(K)$ sur d : il est facile de vérifier en effet que $\lambda(\pi(\sigma, \cdots)) = \dim(\sigma)$.

Pour montrer que la construction est indépendante des choix effectués, on applique encore la proposition 4.5 à $\pi(\sigma, \text{choix1})$ et $\pi(\sigma, \text{choix2})$; l'égalité

$$\langle \pi(\sigma, \text{choix1}), \pi(\sigma, \text{choix2}) \rangle \ =\ \dim.\mathrm{Hom}(\sigma, \sigma) \ =\ 1$$

permet alors de conclure compte tenu de (ii).

Enfin (iii) est une conséquence de la compatibilité des constructions globales à la torsion et au passage aux contragrédientes, aux caractères centraux et aux déterminants ; l'égalité des facteurs locaux résulte des égalités de 4.4.

4.7. Il reste à ce stade à exorciser les deux ennuis restants (d'ailleurs liés entre eux), c'est-à-dire le fait que l'on se soit limité à ne considérer que les σ factorisées à travers un groupe fini fixe, correspondant à une extension E/F, et la dépendance de la construction par rapport au choix d'un corps global $\mathbf{K}$ adapté à cette extension. Ce que prouve Henniart, c'est l'indépendance du résultat pourvu que E *soit assez grand* (une condition qui dépend de σ) : il doit contenir une extension $E(\sigma)$. On s'attend bien sûr à ce que cette condition soit superflue. Il démontre cela par récurrence sur n, en même temps que l'énoncé final (c'est-à-dire le théorème 1).

Cette dernière idée est la suivante : pour σ fixée de degré n, l'égalité des facteurs ϵ, qui entraîne celle des conducteurs, et la connaissance du caractère central, ne laissent qu'un nombre fini de possibilités pour $\pi(\sigma, \mathbf{K})$. Si on avait deux résultats différents $\pi(\sigma, \mathbf{K})$ et $\pi(\sigma, \mathbf{K}')$, il existerait $m < n$ et $\tau \in \mathcal{A}_m^0(K)$ tels que les facteurs ϵ locaux des paires $\pi(\sigma, \mathbf{K}) \times \tau$ et $\pi(\sigma, \mathbf{K}') \times \tau$ soient distincts. Mais par hypothèse de récurrence, τ (dont on peut supposer le caractère central d'ordre fini) est l'image de $\sigma' \in \mathcal{G}_m^0(K)$. La connaissance des facteurs ϵ des paires $\pi(\sigma, \mathbf{K}') \times \pi(\sigma')$ pour un ensemble fini fixé (en fonction de σ) de représentations $\sigma' \in \mathcal{G}_m^0(K)$ (avec des entiers $m < n$) discrimine donc entre les diverses possibilités.

Prenons alors pour $E(\sigma)$ le corps composé du corps fixé par $\mathrm{Ker}(\sigma)$ et des $E(\sigma')$ (supposés définis par récurrence). Si E contient $E(\sigma)$, tout corps global adapté permet aussi de réaliser les $\pi(\sigma')$, ceci sans ambiguïté d'après l'hypothèse de récurrence. Pour deux choix possibles $\mathbf{K}$ et $\mathbf{K}'$, on a, d'après la proposition 4.6 ((iii) :

$$\epsilon(\pi(\sigma), \mathbf{K}') \times \pi(\sigma'), s, \psi) = \epsilon(\sigma \otimes \sigma', s, \psi) = \epsilon(\pi(\sigma), \mathbf{K}') \times \pi(\sigma'), s, \psi);$$

il en résulte que $\pi(\sigma, \mathbf{K}) = \pi(\sigma, \mathbf{K}')$.

Ceci donne finalement une construction 'stable' sans ambiguïté de $\pi(\sigma)$. L'injectivité découle facilement de ce qui précède, puis la surjectivité résulte de la conjecture numérique.

5. OBJETS GÉOMÉTRIQUES LOCAUX : LA THÉORIE DE LUBIN–TATE NON ABÉLIENNE

5.1. On appelle *$\mathcal{O}$–module formel* sur une $\mathcal{O}$–algèbre R un groupe formel Σ de dimension 1, muni d'une action de l'anneau $\mathcal{O}$, de telle sorte que l'action dérivée sur $Lie(\Sigma)$ soit donnée par le morphisme structural $\mathcal{O} \to R$. Dans le cas où R est un corps extension du corps résiduel k de K, on peut définir la *hauteur* h de Σ de la façon suivante : une uniformisante ϖ de K ainsi qu'une coordonnée X sur le groupe formel étant choisies, q^h est alors l'ordre de la série formelle en X qui exprime l'action de ϖ (dans la suite cet ordre et donc la hauteur seront toujours finis). On ne confondra pas cette notion avec la hauteur de Σ vu simplement comme groupe formel (dans le cas où K est un corps p–adique, la hauteur du groupe est égale à $h.\deg[K : \mathbf{Q}_p]$).

C'est Drinfeld qui le premier a étudié ces objets. Ils constituent une généralisation facile de la théorie usuelle des groupes formels (théorie que l'on retrouve dans le cas particulier où $K = \mathbf{Q}_p$). Les résultats habituels se transposent sans mal : en particulier, il existe sur $\overline{k}$, clôture algébrique du corps résiduel de K, un module formel Σ_h de hauteur h unique à isomorphisme près. L'anneau des endomorphismes de Σ_h est isomorphe à l'ordre maximal du corps gauche $D_{K,h}$ de centre K de dimension h^2 et d'invariant $1/h$ dans le groupe de Brauer.

On s'intéresse à la théorie des *déformations* de Σ_h : c'est-à-dire au foncteur Def qui à une $\mathcal{O}$–algèbre artinienne R de corps résiduel $\overline{k}$ associe l'ensemble des classes d'isomorphisme de déformations de Σ_h sur R ($\mathcal{O}$–modules formels sur R munis d'un isomorphisme entre leur réductions modulo l'idéal maximal de R et Σ_h). Comme Lubin l'avait fait pour les groupes formels, on montre que Def est (pro-)représentable par une algèbre locale complète $R_{K,h}$, en fait isomorphe simplement à l'algèbre des séries formelles en $h-1$ indéterminées sur $\hat{\mathcal{O}}^{\mathrm{nr}}$, anneau des entiers du complété d'une extension non-ramifiée maximale K^{nr} de K. Sur le spectre formel de $R_{K,h}$, on a donc la déformation universelle $\tilde{\Sigma}_h$ de Σ_h.

5.2. Soit $m \geq 1$ un entier (le niveau). Au-dessus de la fibre générique de $\mathrm{Spf}(R_{K,h})$, le sous-groupe de ϖ^m–torsion $(\tilde{\Sigma}_h)_{\varpi^m}$ est étale localement isomorphe à $(\varpi^{-m}\mathcal{O}/\mathcal{O})^h$. Si x est un point de cette fibre générique, la donnée d'une *structure de niveau* m en ce point consiste en la donnée d'un isomorphisme

$$\phi : (\varpi^{-m}\mathcal{O}/\mathcal{O})^h \to (\tilde{\Sigma}_h)_{\varpi^m}(x).$$

Drinfeld a étendu cette notion de structure de niveau au cas d'un point quelconque, où le groupe $(\tilde{\Sigma}_h)_{\varpi^m}$ n'est donc plus nécessairement étale : c'est la notion de *'base de Drinfeld'*, pour laquelle on renvoie à [Dr1] et [K-M].

Le foncteur qui à une $R_{K,h}$–algèbre artinienne R associe l'ensemble des bases de Drinfeld de niveau m sur $(\tilde{\Sigma}_h)|_{\mathrm{Spec}R}$ est pro-représentable par un anneau local complet $R_{K,h,m}$. D'autre part, pour $m_1 \geq m_2$, une base de Drinfeld de niveau m_1 se restreint en une base de niveau m_2 : cela est clair au moins au niveau des fibres génériques. Les schémas formels $\mathrm{Spf}(R_{K,h,m})$ constituent donc, m variant, un système projectif. En particulier, l'oubli de la structure de niveau définit une projection :

$$\mathrm{Spf}(R_{K,h,m}) \to \mathrm{Spf}(R_{K,h,0}) = \mathrm{Spf}(R_{K,h}).$$

Enfin on a une action du groupe $GL_h(\mathcal{O}/\varpi^m\mathcal{O})$ sur $\mathrm{Spf}(R_{K,h,m})$, compatible à la projection ci-dessus. Cette projection est ainsi un revêtement galoisien de groupe $GL_h(\mathcal{O}/\varpi^m\mathcal{O})$, étale en fibre générique, mais au contraire très ramifié au-desssus de la fibre spéciale.

5.3. Les objets locaux fondamentaux sont les espaces de *cycles évanescents* ℓ–adiques (avec $\ell \neq p$ un nombre premier fixé), notés $\Psi^i_{K,\ell,h,m}$, associés par la théorie de Berkovich (cf. [Be1] [Be2] [Be3]) aux schémas formels $\mathrm{Spf}(R_{K,h,m})$. La définition de Berkovich utilise la théorie des espaces analytiques p–adiques (la fibre générale d'un schéma formel appartient à cette catégorie). Lorsque K et ℓ seront fixés, nous les éliminerons de la notation pour alléger. Ces espaces constituent pour m variant un système inductif (à morphismes injectifs), dont nous noterons $\Psi^i_{K,\ell,h}$ ou plutôt Ψ^i_h la limite (réunion).

5.4. On dispose sur les $\Psi^i_{h,m}$, et donc sur leur limite, d'actions (qui commutent entre elles) des trois groupes suivants :

(i) Le groupe $GL_h(\mathcal{O})$ qui agit via son quotient $GL_h(\mathcal{O}/\varpi^m\mathcal{O})$ comme expliqué ci-dessus (action sur la structure de niveau).

(ii) Le groupe $\mathcal{O}^*_{D_h}$ des éléments inversibles de l'ordre maximal $\mathcal{O}_{D_h}$ du corps gauche $D_h = D_{K,h}$: en effet ce groupe est le groupe des automorphismes de Σ_h et en tant que tel il agit sur l'espace de la déformation universelle et sur $\tilde{\Sigma}_h$.

(iii) Enfin le groupe d'inertie $\mathrm{Gal}(\overline{K}/K^{\mathrm{nr}})$.

On renvoie le lecteur à [Ca3] ou [H-T] pour une explication du fait suivant : le produit des trois actions considérées peut s'étendre en une action sur la limite Ψ^i_h du sous-groupe $A_h \subset GL_h(K) \times D_h^* \times W_K$ constitué des triples (g, d, w) qui vérifient la

relation : $det(g)^{-1}N(d)\mathrm{Cl}(w) \in \mathcal{O}^*$, où N désigne la norme réduite et Cl : $W_K \to K^*$ l'application de réciprocité. En utilisant la théorie de Berkovich, on peut montrer que l'action de A_h est admissible.

5.5. Soit τ une représentation admissible irréductible de D_h^*. On note

$$\Psi_h^i(\tau) = \mathrm{Hom}_{\mathcal{O}_{D_h}^*}(\tau, \Psi_h^i),$$

la composante isotypique pour la restriction de τ à $\mathcal{O}_{D_h}^*$. On obtient ainsi une représentation du produit $GL_h(K) \times W_K$: un couple (g, w) étant donné, on le complète en un triple $(g, d, w) \in A_h$ et on définit comme suit l'action de (g, w) sur $\phi \in \Psi_h^i(\tau)$:

$$(g, w).\phi(t) = (g, d, w).\phi(d^{-1}.t).$$

Si l'on préfère, on peut aussi considérer la représentation induite $\mathrm{Ind}_{A_h}^{GL_h(K)\times D_h^*\times W_K}(\Psi_h^i)$ et alors $\Psi_h^i(\tau)$ s'identifie à la composante τ–isotypique dans cette induite.

Soit τ choisie de telle sorte que $\pi = \mathrm{JL}(\tau)$ soit cuspidale (cf. (3.7) ci-dessus). La conjecture énoncée dans [Ca3] prédit que l'on doit avoir alors une équivalence :

$$\Psi_h^{h-1}(\tau) = \mathrm{JL}(\tau)^\vee \otimes \left(\sigma_h(\mathrm{JL}(\tau) \otimes |\ |^{\frac{1-h}{2}}\right).$$

D'autre part, toujours sous l'hypothèse que $\pi = \mathrm{JL}(\tau)$ est cuspidale, on s'attend à ce que les $\Psi_h^i(\tau)$ soient nuls pour $i \neq h-1$. Lorsque $h = 1$, vérifier tout cela constitue un exercice facile de traduction de la théorie de Lubin–Tate (en particulier, $R_{K,1}$ est alors isomorphe à $\mathcal{O}^{\mathrm{nr}}$, $(\tilde{\Psi}_h)$ à un groupe de Lubin–Tate et enfin $R_{K,1,m}$ à l'anneau des entiers de l'extension abélienne de K^{nr} engendrée par les points de ϖ^m–torsion du Lubin–Tate). Lorsque $h = 2$ et que K est un corps p–adique, j'avais vérifié cette conjecture ([Ca2]) en généralisant certaines idées que Deligne avait développées dans le cas du corps $\mathbf{Q}_p$ et en utilisant ensuite le changement de base automorphe.

Harris et Taylor considèrent plutôt la représentation virtuelle dans le groupe de Grothendieck des représentations admissibles de longueur finie de $GL_h(K) \times W_K$:

$$\Psi_h^*(\tau) = (-1)^{h-1} \sum (-1)^i [\Psi_h^i(\tau)].$$

Une des étapes importantes de leur article consiste à démontrer le théorème suivant (pour K un corps p–adique) :

5.6. THÉORÈME 2.— *Soit τ comme ci-dessus, telle que $\pi = \mathrm{JL}(\tau)^\vee$ soit cuspidale. Alors il existe une 'vraie' représentation*

$$r_h(\pi) : W_K \to GL_h(\overline{\mathbf{Q}}_\ell)$$

telle que l'on ait l'égalité dans le groupe de Grothendieck :

$$\Psi_h^*(\tau) = [\pi \otimes r_h(\pi)].$$

On définira alors

$$\sigma_h(\pi^\vee) = r_h(\pi) \otimes |\ |^{\frac{h-1}{2}}$$

et il 'reste' alors à prouver que cela possède les propriétés que l'on attend de la correspondance de Langlands locale. Je crois qu'il est assez facile ensuite d'en déduire la conjecture telle que je l'avais énoncée (pour Ψ_h^{h-1} plutôt que pour Ψ_h^*).

5.7. Bien qu'ils n'interviennent pas directement dans l'article de Harris et Taylor, on a aussi étudié une autre sorte d'objets locaux : les espaces habituellement notés Ω_K^h sont des espaces analytiques rigides sur le corps local K qui constituent pour $h = 2$ un analogue p–adique du demi-plan de Poincaré (et une généralisation en dimension supérieure si $h > 2$). L'ensemble des points sur $\widehat{\overline{K}}$ de Ω_K^h est le complémentaire dans l'espace projectif $\mathbf{P}^{h-1}(\widehat{\overline{K}})$ de la réunion des hyperplans K–rationnels. On voit (au moins ensemblistement) que Ω_K^h est muni d'une action du groupe $GL_n(K)$.

Drinfeld ([Dr2]) a interprété Ω_K^h comme classifiant d'une certaine famille de groupes formels et en a déduit la construction d'un système projectif de revêtements rigides analytiques $GL_n(K)$–équivariants :

$$\mathrm{Rev}_K^{h,m} \to \Omega_K^h \otimes_F \hat{F}^{\mathrm{nr}}$$

galoisiens de groupes $[\mathcal{O}_{D_h}/\varpi^m\mathcal{O}_{D_h}^*]^*$. La situation est en quelque sorte inversée par rapport à nos revêtements précédents $\mathrm{Spf}(R_{K,h,m}) \to \mathrm{Spf}(R_{K,h})$ où c'était $\mathcal{O}_{D_h}^*$ qui opérait sur la base, et $GL_h(\mathcal{O}/\varpi^m\mathcal{O})$ le groupe de Galois.

On peut considérer la cohomologie ℓ–adique à support propre (définie par Berkovich) de ces revêtements :

$$\Psi'^i_{h,m} = H^i_{\mathrm{c}}(\mathrm{Rev}_K^{h,m}, \overline{\mathbf{Q}}_\ell)$$

ainsi que leur limite Ψ'^i_h. On vérifie que c'est un espace où opère le sous–groupe $A'_h \subset GL_h(K) \times D_h^* \times W_K$ constitué des triples (g, d, w) qui vérifient la relation :

$det(g)N(d)\mathrm{Cl}(w)^{-1} \in \mathcal{O}^*$. On considère alors l'induite de A'_h à $GL_h(K) \times D_h^* \times W_K$ de Ψ'^i_h, puis, pour π une représentation cuspidale de $GL_h(K)$, la composante π–isotypique dans cette induite que nous noterons $\Psi'^i_h(\pi)$. J'avais conjecturé dans le même article ([Ca3]) que l'on devait avoir

$$\Psi'^{h-1}_h(\pi) = \mathrm{JL}(\pi) \otimes (\sigma_h(\pi^\vee) \otimes |\ |^{\frac{1-h}{2}}) .$$

Dans [Ha1], Harris avait prouvé partiellement (une variante de) cette conjecture. C'est-à-dire qu'il ne déterminait pas complètement dans tous les cas la représentation galoisienne associée à π. Combiné maintenant avec le présent travail, cela permet de traiter tous les cas et de prouver cette conjecture complètement.

5.8. Pour résumer les choses : l'une et l'autre des deux représentations locales ainsi définies Ψ^{h-1}_h et Ψ'^{h-1}_h doivent constituer une construction géométrique simultannée des correspondances de Langlands et de Jacquet–Langlands locales pour les cuspidales. Noter toutefois la légère différence entre la formule donnant $\Psi^{h-1}_h(\pi)$ et celle exprimant $\Psi'^{h-1}_h(\pi)$. Ce qu'il advient des autres représentations (en particulier des séries discrètes non cuspidales) n'est pas encore clair. Voir toutefois [S-S] pour un calcul complet de la cohomologie de Ω^h_K.

Pour $h = 1$, tout cela n'est que la théorie de Lubin–Tate. À partir de $h = 2$, la plupart des résultats s'obtiennent par des méthodes globales ; voir toutefois Faltings [Fa] pour une étude purement locale de la représentation virtuelle Ψ'^*_h somme alternée des Ψ'^i_h (mais sans l'action galoisienne).

Dans le cas des corps locaux d'égale caractéristique p, Boyer [Boy] vient de démontrer dans sa thèse la conjecture pour la représentation Ψ^{h-1}_h. On peut espérer faire de même pour la représentation Ψ'^{h-1}_h (travail en cours d'un de mes étudiants).

La relation entre les représentations Ψ^i_h et Ψ'^i_h constitue un autre mystère. Tant les espaces définis en 3.2 que les espaces de Drinfeld sont d'ailleurs des cas particuliers d'objets locaux construits par Rapoport et Zink et destinés à décrire localement les variétés de Shimura (voir [R-Z] et [Bo]).

6. VARIÉTÉS DE SHIMURA

6.1. On fixe un corps de nombres F de degré $2d$, composé d'un corps totalement réel F^+ de degré d et d'un corps quadratique imaginaire E. Le lien avec le corps local K précédemment considéré est le suivant : K s'identifie au complété de F^+ en une

place (au-dessus de p) notée $v = v_1$. On désigne par $v_2, \cdots v_r$ les autres places de F^+ divisant p. D'autre part on suppose que p se décompose dans E en deux places $\mathfrak{p}$ et $\overline{\mathfrak{p}}$. Les places de F divisant p sont notées $w = w_1, w_2, \cdots w_r$ (au-dessus de $\mathfrak{p}$ et de $v_1, v_2, \cdots v_r$) et $\overline{w} = \overline{w}_1, \overline{w}_2, \cdots \overline{w}_r$ (leurs conjuguées complexes), de sorte que K s'identifie aussi au complété de F en w ou en $\overline{w}$.

B désigne une algèbre à division de centre F et de dimension n^2 sur F munie d'une involution de seconde espèce positive $*$. On fait aussi quelques hypothèses techniques (susceptibles d'évoluer dans les versions successives de l'article [H-T]) : le fait qu'en chaque place B soit déployée ou soit un corps gauche ; cela est rendu nécessaire par l'état actuel de la théorie de la correspondance de Jacquet–Langlands globale. Conjuguant l'involution $*$ par un élément $*$–antisymétrique $\beta \in B$, on obtient une autre involution $\sharp$.

Intéressons-nous au groupe associé à l'involution $\sharp$: c'est un groupe réductif G sur $\mathbf{Q}$ dont l'ensemble des points à valeurs dans une $\mathbf{Q}$–algèbre R est donné par :

$$G(R) = \{g \in B^{op} \otimes R \ ; \ bb^\sharp \in R^*\}.$$

C'est une forme du groupe des similitudes unitaires de degré n associé au corps F. On fixe β de telle sorte *qu'à l'infini ce groupe soit de type* $(n-1, 1)$ *en une place* τ_1 *de* F^+ *et* $(n, 0)$ *aux autres.* L'autre hypothèse fondamentale est que B *est déployée en* w. On fixe un isomorphisme : $B \otimes F_w \simeq M_n(F_w)$.

On considère alors les variétés de Shimura associées au groupe G. Elles dépendent pour leur définition du choix d'une classe de conjugaison d'homomorphismes $h : \mathbf{C}^* \to G(\mathbf{R})$ (voir Deligne [De1] pour le formalisme) et d'un sous–groupe compact ouvert $U \subset G(\mathbf{A}^\infty)$. Ce sont des variétés algébriques projectives S_U définies sur le corps F, dont l'ensemble des points complexes ($\mathbf{C}$ étant vu comme une $\mathbf{Q}$–algèbre par le plongement τ_1 combiné avec un plongement complexe de E) est donné par :

$$S_U(\mathbf{C}) = G(\mathbf{Q}) \setminus [\mathcal{X} \times G(\mathbf{A}^\infty)/U],$$

avec $\mathcal{X}$ l'espace symétrique associé à $G(\mathbf{R})$ (c'est-à-dire la boule unité de $\mathbf{C}^{n-1}$). Plus concrètement, c'est une réunion de quotients $\Gamma \setminus \mathcal{X}$ de $\mathcal{X}$ par des groupes de congruence Γ de $G(\mathbf{R})$.

Sous les hypothèses qui précèdent (essentiellement le type archimédien de G et son déploiement en w), on peut décrire la réduction de la variété S_U à la place w, par une théorie qui généralise en quelque sorte la théorie des courbes modulaires. L'objet de ce paragraphe est d'expliquer comment.

Pour alléger, nous noterons désormais simplement S_U la variété $S_U \otimes F_w$ sur $F_w \simeq K$.

6.2. On sait d'après Shimura et Deligne que S_U représente un foncteur défini en termes de variétés abéliennes : il s'agit de classifier des variétés abéliennes A à isogénie près de dimension dn^2, munies d'une action de B et d'une polarisation, telles que l'involution de Rosati associée induise l'involution $*$ sur B ; une certaine condition décrivant l'algèbre de Lie de A doit être satisfaite et on ajoute en plus une structure de niveau que nous détaillerons ci-dessous. À vrai dire, on obtient ainsi un problème de modules représentable par S_U seulement dans le cas où n est pair, tandis que dans le cas où n est impair on obtient une somme X_U de copies de S_U. Je décrirai ci-dessous X_U sans trop me soucier de ces questions liées à des obstructions au principe de Hasse (pour des tores).

Le point central qui permet de décrire la réduction (éventuellement mauvaise) des variétés étudiées tient en la forme particulière que prend dans ce cas la condition sur $Lie(A)$, pour une variété abélienne sur une F_w-algèbre R. On utilise la décomposition :

$$F \otimes \mathbf{Q}_p = F_{w_1} \oplus F_{w_2} \oplus \cdots \oplus F_{w_r} \oplus F_{\overline{w}_1} \oplus F_{\overline{w}_2} \oplus \cdots \oplus F_{\overline{w}_r},$$

ce qui induit une décomposition semblable sur tout $F \otimes \mathbf{Q}_p$-module L et le décompose en sous-modules que nous noterons L_{w_i} et $L_{\overline{w}_i}$. En particulier, tout $B \otimes \mathbf{Q}_p$ module L se décompose en la somme de B_{w_i}-modules L_{w_i} et de $B_{\overline{w}_i}$-modules $L_{\overline{w}_i}$. On peut aller plus loin dans la décomposition du $M_n(F_w)$-module L_{w_1}, qui définit un F_w-module eL_{w_1} par l'action de l'idempotent

$$e = \begin{pmatrix} 1 & 0 & 0 & 0 \cdots & 0 \\ 0 & 0 & 0 & 0 \cdots & 0 \\ 0 & 0 & 0 & 0 \cdots & 0 \\ . & . & . & . \cdots & . \\ 0 & 0 & 0 & 0 \cdots & 0 \end{pmatrix} ;$$

réciproquement, on récupère L_{w_1} comme $M_n(eL_{w_1})$.

On peut appliquer cela à l'algèbre $Lie(A)$ pour A une variété abélienne sur une F_w-algèbre R, munie d'une action de B. Cette algèbre de Lie est un R-module projectif muni d'une action de $B \otimes \mathbf{Q}_p$, et donc par restriction de $F \otimes \mathbf{Q}_p$. La condition mentionnée ci-dessus est alors que :

(i) pour $i \geq 2$: $Lie(A)_{w_i}$ est nul.

(ii) $eLie(A)_w$ est un R–module (projectif) de rang 1, l'action de F_w qui provient de la décomposition coïncidant avec celle qui provient de l'inclusion $F_w \hookrightarrow R$.

6.3. Supposons le sous-groupe U décomposé en un produit $U^p \times U_p$, avec $U^p \subset G(\mathbf{A}^{p,\infty})$ (adèles finies sans la composante en p) et

$$U_p \subset G(\mathbf{Q}_p) \simeq \mathbf{Q}_p^* \times GL_n(F_{w_1}) \times B_{w_2}^* \times \cdots \times B_{w_r}^* ,$$

que l'on supposera également décomposé en le produit du sous-groupe $\mathbf{Z}_p^*$ de $\mathbf{Q}_p^*$ et de sous-groupes $U_{w_1} \subset GL_n(F_{w_1})$ et $U_{w_i} \subset B_{w_i}^*$.

D'autre part, V désigne l'espace vectoriel B, muni de la forme alternée $(b_1, b_2) = \mathrm{tr}_{B/\mathbf{Q}}(b_1 \beta b_2^*)$. Pour A une variété abélienne (en caractéristique 0) polarisée et munie d'une action compatible de B, une *structure de niveau* U est la donnée, localement pour la topologie étale, d'une classe modulo U d'isomorphismes symplectiques B–linéaires $\prod V_l(A) \simeq V \otimes \mathbf{A}^\infty$, avec V_l le module de Tate tensorisé par $\mathbf{Q}$.

Fixons quelques structures entières : soit $\mathcal{O}_{F(p)}$ l'ensemble des éléments de F qui sont entiers en p ; on choisit un $\mathcal{O}_{F(p)}$–ordre $\mathcal{O}_B \subset B$ stable par l'involution $*$, tel que sa complétion en p soit maximale et tel que $\mathcal{O}_{B,w_1} = M_n(\mathcal{O}_{F,w_1})$. On choisit aussi un réseau $\Lambda \subset V$ stable par $\mathcal{O}_B$, tel que $\Lambda_{w_i} = \mathcal{O}_{Bw_i}$ et tel que la forme $(\ ,\ )$ induise une autodualité parfaite (à valeurs dans $\mathbf{Z}_p$) sur $\Lambda \otimes \mathbf{Z}_p$. On suppose enfin que les U_{w_i} sont inclus dans les $\mathcal{O}_{B,w_i}^*$.

Plaçons-nous dans un premier temps sous l'hypothèse que $U_{w_1} = U_w$ *coïncide avec le sous-groupe compact maximal* $GL_n(\mathcal{O}_{F,w}) \subset GL_n(F_w)$. On montre alors que X_U a *bonne réduction.* Plus précisément, le foncteur $\mathcal{F}$ qui suit, défini sur les $\mathcal{O}_{F,w} \simeq \mathcal{O}$–algèbres, est représentable par un $\mathcal{O}$–schéma projectif lisse de fibre générique X_U. Pour alléger, nous continuerons à noter X_U le $\mathcal{O}$–schéma ainsi obtenu.

$\mathcal{F}(R)$ est l'ensemble des classes de $(r+3)$–uples $(A, \lambda, \iota, \bar{\eta}^p, \bar{\phi}_2, \cdots \bar{\phi}_r)$ avec :

- *A un schéma abélien sur R de dimension relative dn^2.*
- *λ une polarisation de A.*
- *$\iota : \mathcal{O}_B \hookrightarrow \mathrm{End}(A) \otimes \mathbf{Z}_{(p)}$ une action.*
- *On impose que : $\lambda \circ \iota(b) = \iota(b^*)^\vee \circ \lambda$ et les conditions sur $Lie(A)$ analogues à* (i) *et* (ii) *ci-dessus. Noter en effet que l'on a une décomposition de $\mathcal{O}_{F(p)} \otimes \mathbf{Z}_p$ analogue à la décomposition précédente de $F \otimes \mathbf{Q}_p$, de sorte que la décomposition ci-dessus de $Lie(A)$ vaut aussi sur les entiers.*

• $\bar{\eta}^p$ *est la donnée, localement pour la topologie étale, d'une classe modulo* U^p *de similitudes symplectiques* $\mathcal{O}_B$*–linéaires :*

$$\eta^p : \prod_{l \neq p} V_l(A) \simeq V \otimes \mathbf{A}^{p,\infty}.$$

• $\bar{\phi}_j$ *(pour* $2 \leq j \leq r$*) la donnée, localement pour la topologie étale, d'une classe modulo* U_{w_i} *d'isomorphismes (avec des notations que nous clarifierons ci-dessous) :*

$$\phi_j : T_p(A)_{w_i} \simeq \Lambda_{w_i}.$$

On identifie deux tels $r+3$*–uples* $(A, \lambda, \iota, \bar{\eta}^p, \bar{\phi}_2, \cdots \bar{\phi}_r)$ *et* $(A', \lambda', \iota', \bar{(\eta^p)}', \bar{\phi}'_2, \cdots \bar{\phi}'_r)$ *s'il existe une isogénie d'ordre premier à* p *qui échange toutes les structures (seulement à un élément de* $\mathbf{Z}^*_{(p)}$ *près en ce qui concerne la polarisation).*

6.4. Le point crucial de toute la théorie de la réduction de notre variété à la place w est le suivant : on a pour le groupe p–divisible A_{p^∞} associé à A une décomposition :

$$A_{p^\infty} = A_{w_1^\infty} \oplus A_{w_2^\infty} \oplus \cdots \oplus A_{w_r^\infty} \oplus A_{\overline{w}_1^\infty} \oplus A_{\overline{w}_2^\infty} \oplus \cdots \oplus A_{\overline{w}_r^\infty}$$

où l'on a noté $A_{w_i^\infty}$ et $A_{\overline{w}_i^\infty}$ les composants de la décomposition (comme ci–dessus) du groupe sous l'action de $\mathcal{O}_{F(p)} \otimes \mathbf{Z}_p \subset \mathcal{O}_B \otimes \mathbf{Z}_p$. Les conditions imposées sur l'algèbre de Lie forcent les groupes $A_{w_2^\infty} \cdots A_{w_r^\infty}$ à être ind-étales (ce qui fait que les structures de niveau sont les 'mêmes' qu'en caractéristique 0 : on peut en particulier définir (pour $i \geq 2$) les modules de Tate partiels $T_p(A)_{w_i}$ déjà utilisés ci-dessus). D'autre part, la partie $eA_{w_1^\infty}$ de $A_{w_1^\infty}$ doit être un groupe p–divisible $\mathcal{G}_A$ muni d'une action de $\mathcal{O}_{F,w} = \mathcal{O}$, dont la partie formelle $\mathcal{G}_A^0$ (on se place ici dans le cas où R est un anneau local artinien de caractéristique résiduelle p) doit être de dimension 1, l'action dérivée se factorisant à travers le morphisme structural. C'est ce genre d'objets que Drinfeld nomme '$\mathcal{O}$–modules divisibles' ; la partie formelle est alors un $\mathcal{O}$–module formel au sens du paragraphe 5. Enfin, pour R un corps de caractéristique p, la hauteur totale de $\mathcal{G}_A$ est égale à n : cette hauteur se répartit entre la partie formelle, de hauteur $n-h$, et le quotient étale $\mathcal{G}_A^{\text{et}}$ de hauteur h (c'est-à-dire localement isomorphe à $(K/\mathcal{O})^h$) ; ici h est un entier compris entre 0 et $n-1$.

La remarque essentielle est qu'il 'revient au même' de déformer $(A, \lambda, \iota, \bar{\eta}^p, \bar{\phi}_2, \cdots \bar{\phi}_r)$ ou de déformer $\mathcal{G}$ (avec sa structure de $\mathcal{O}$–module) : en effet, d'après le théorème de Serre et Tate, déformer A revient à déformer son groupe p–divisible (muni des différentes structures déduites de celles de A). Mais les $A_{w_j^\infty}$ pour $j \geq 2$

étant étales, leurs déformations sont triviales. Une fois $\mathcal{G}$ déformé, on obtient une déformation de $A_{w_1^\infty} = M_n(\mathcal{G})$ et donc une de $\oplus_{j=1}^r A_{w_j^\infty}$. D'autre part, les conditions imposées font que la polarisation met les deux groupes p–divisibles $\oplus_{j=1}^r A_{w_j^\infty}$ et $\oplus_{j=1}^r A_{\overline{w}_j^\infty}$ en dualité de Cartier. Si l'on a déformé la première moitié, on récupère donc automatiquement une déformation de la seconde.

Le '$\mathcal{O}$–module divisible' $\mathcal{G}_A$ gouverne donc la théorie des déformations de A. C'est cela qui permet d'étudier localement le problème de modules ci-dessus. C'est aussi ce qui permet de comparer la structure locale de X_U au voisinage d'un point aux espaces locaux considérés au paragraphe précédent. Ces idées jouaient déjà un rôle crucial dans mon étude [Ca1] des courbes de Shimura, et elles sont ici généralisées à des variétés de dimension supérieure.

Ceci étant, on peut alors introduire des structures de niveau en w : supposons que $U_{w_1} = U_w$ est maintenant le sous–groupe de congruence de niveau m de $GL_n(O_{F,w})$, constitué des matrices congrues à 1 modulo ϖ^m. Dans ce cas, on considère le problème de modules obtenu en rajoutant au précédent la structure supplémentaire suivante :

- $\phi_1 : (\varpi^{-m}\mathcal{O}/\mathcal{O})^n \to (\mathcal{G}_A)_{\varpi^m}(R)$, *une base de Drinfeld de niveau m (c'est-à-dire dont l'image en tant que sous-schéma coïncide avec* $(\mathcal{G}_A)_{\varpi^m}|_R$*).*

Ce problème de modules est alors représentable par un $\mathcal{O}$–schéma $X_U = X_{U^w,m}$ projectif (mais non plus lisse) (où l'on note U^w le produit de U^p et des U_{w_j} pour $j \geq 2$).

6.5. La fibre spéciale $\overline{X}_{U^w,m}$ de $X_U = X_{U^w,m}$ admet une filtration par des sous-variétés fermées $\overline{X}^{[h]}_{U^w,m}$: le lieu où la hauteur de la partie étale est $\leq h$. On note $\overline{X}^{(h)}_{U^w,m} = \overline{X}^{[h]}_{U^w,m} - \overline{X}^{[h-1]}_{U^w,m}$. Ce sont des sous-ensembles localement fermés de dimension h, lisses si $m = 0$. En particulier, les *'points supersinguliers'* sont ceux pour lesquels $\mathcal{G}_A$ est purement infinitésimal et ils constituent un ensemble fini $\overline{X}^{[0]}_{U^w,m}$. D'après ce qui a été dit plus haut, le complété de l'anneau local de $X_{U^w,m}$ en un tel point géométrique de la fibre spéciale est isomorphe à l'anneau $R_{K,n,m}$ de 5.2. Les cycles évanescents formels (mais on sait d'après Berkovich qu'ils coïncident avec les cycles évanescents usuels de la géométrie algébrique) sont donc exprimables en termes des espaces $\Psi^i_{K,n,m}$ de 5.3.

7. LE PRINCIPE DE LA MÉTHODE D'HARRIS ET TAYLOR

7.1. Soit Π une représentation automorphe du groupe $G(\mathbf{A})$. Lorsque la composante archimédienne Π_∞ de Π est une certaine représentation de la série discrète, Π 'apparaît' dans la cohomologie $H^{n-1}(X_{U^w,m} \times \overline{F}, \mathcal{L}_\xi)$, où $\mathcal{L}_\xi$ est un $\overline{\mathbf{Q}}_\ell$–faisceau lisse associé à une représentation algébrique du groupe G. La composante $G(\mathbf{A}^\infty)$–isotypique correspondant à Π^∞ dans

$$\varinjlim_{U^w,m} H^{n-1}(X_{U^w,m} \times \overline{F}\ ,\ \mathcal{L}_\xi)$$

est alors une représentation ℓ–adique du groupe $\mathrm{Gal}(\overline{F}/F)$. C'est à peu près ainsi que l'on définit les représentations $\rho_{\lambda,\Pi}$ dont il a été question plus haut (cf. 3.3), à deux détails près : d'une part, dans la notation de 3.3, Π désignait la représentation automorphe de $GL_n(\mathbf{A}_F)$ correspondant par changement de base à notre présente Π ; d'autre part, ainsi qu'on l'a déja remarqué, on récupère ainsi (après semi-simplification) un multiple de $\rho_{\lambda,\Pi}$ par un entier a dont on ne sait pas montrer qu'il est égal à 1.

Aux bonnes places v, on montre que $(\rho_{\lambda,\Pi})_v$ et Π_v se correspondent par la correspondance 'de Hecke' locale (2.2). Cela provient d'une comparaison entre des formules des traces de Lefschetz et de Selberg, en utilisant les résultats de Kottwitz, lesquels contiennent une description de l'ensemble des points de la fibre spéciale (la variété considérée a alors bonne réduction).

Pour aller plus loin, c'est-à-dire obtenir un résultat analogue aux mauvaises places, il est naturel de faire usage de la suite spectrale des cycles évanescents pour $X_{U^w,m}/\mathcal{O}$:

$$H^i(\overline{X} \otimes_k \overline{k}\ ,\ R\Phi^j \otimes \mathcal{L}_\xi) \Longrightarrow H^{i+j}(X \times \overline{F}_w\ ,\ \mathcal{L}_\xi),$$

où $\overline{X}$ désigne la fibre spéciale de $X = X_{U^w,m}$ et $R\Phi^j$ les faisceaux de cycles évanescents ; en fait $R\Phi^j$ est concentré sur la strate fermée $\overline{X}^{[n-1-j]}$. Il est facile de voir (supposant bien sûr que $\ell \neq p$) que les faisceaux $\mathcal{L}_\xi$ sont lisses sur le $\mathcal{O}$–schéma X tout entier, de sorte que le calcul des cycles évanescents se ramène au cas du faisceau constant, ce qui justifie les notations que nous avons employées.

7.2. On a déja observé plus haut le lien aux points supersinguliers entre les groupes de cycles évanescents Φ^j et les espaces $\Psi^j_{K,n,m}$ définis à partir des objets locaux du paragraphe 5. L'idée originelle de Deligne ([De3]) concernait les points supersinguliers des courbes modulaires. Il calculait le groupe de sections (sur la strate supersingulière) $H^0(\overline{X}, R\Phi^1)$ en termes de la représentation locale Ψ^1_2 de 5.3 et obtenait de la sorte un

avatar du théorème 2 ci-dessus (dans le cas $h = 2$). Mais il ne savait montrer que la représentation galoisienne obtenue $r_h(\pi)$ était bien donnée par la correpondance 'de Hecke' que pour π induite automorphe d'un caractère multiplicatif d'une extension quadratique (on disait alors 'ordinaire'). L'argument de Deligne avait été complété par un lemme de Brylinski [Br] – qui remplaçait pour les courbes les résultats de Berkovich dont on a besoin en dimension supérieure. Je l'avais généralisé aux courbes de Shimura, ce qui donnait l'analogue du th. 2 pour $h = 2$ et K un corps p–adique quelconque, puis, par des arguments de changement de base, le fait que $\pi \to r_2(\pi)$ était bien la correspondance voulue.

Je crois que la méthode de Deligne peut s'appliquer telle quelle aux variétés que nous considérons ici. J'avais suggéré cette possibilité dans [Ca3]. Personne ne l'a écrite pour les corps p–adiques dans le cas $n > 2$; comme on le verra, Harris et Taylor obtiennent un résultat plus général par une méthode qui consiste à considérer toutes les strates à la fois et qui repose sur un calcul de traces (alors que Deligne avait un argument direct). En égale caractéristique p, Boyer [Boy] a développé l'idée (pour n quelconque) d'une façon assez proche de sa forme initiale. Son application ici donnerait un analogue du théorème 2 : on obtiendrait de la sorte une application qui à $\pi \in \mathcal{A}_n^0$ associe une représentation galoisienne de dimension n : $r_n(\pi)$ satisfaisant au th. 2 et 'calculant' localement la correspondance globale en le sens suivant : si à une représentation automorphe Π, de composante locale Π_w *cuspidale*, correspond la représentation ℓ–adique Σ , alors la restriction Σ_w de Σ au groupe de galois local est donnée par : $\Sigma_w = r_n(\Pi_w)$ Cette relation entre le local et le global force alors, ainsi qu'on le voit sans peine, la correspondance ainsi obtenue (après torsion idoine et passage à la contragrédiente) $\pi \to \sigma_n(\pi)$ à être compatible avec toutes les opérations et fonctorialités locales qui admettent une extension globale (changement de base et induction automorphe, torsion par un caractère et passage aux contragrédientes). De cela il résulte alors par des arguments dus à Henniart que c'est une bijection, puis d'après Harris ([Ha2]) que, lorsque $n < p$, c'est bien la correspondance voulue.

À vrai dire Harris étudiait dans ce dernier article la même correspondance (du moins rétrospectivement !), mais définie autrement : à partir des espaces de Drinfeld dont nous avons un peu parlé ci-dessus (3.7), autrement dit avec la représentation locale ${\Psi'}_n^{n-1}$ remplaçant Ψ_n^{n-1}. La méthode d'étude est formellement analogue à celle de Deligne, à ceci près que l'objet global que l'on considère dans ce cas est une variété de Shimura analogue à celles considérées ci-dessus, mais provenant d'une algèbre B dont la complétée en $w = w_1$ est maintenant *un corps gauche d'invariant* $1/n$. La relation entre le local et le global transite alors par un *théorème d'uniformisation*

p–adique (Rapoport, Rapoport–Zink, Varshavsky). Ce résultat (qui généralise celui de Čerednik pour les courbes) affirme que les variétés de Shimura considérées sont analytiquement isomorphes (en tant qu'espaces rigides p–adiques) à des réunions de quotients des espaces $\mathrm{Rev}_K^{n,m}$ de 3.7 par des sous-groupes discrets de $GL_n(K)$.

Par l'une ou l'autre de ces méthodes, le résultat obtenu n'est toutefois pas assez fort pour entraîner le théorème 1 dans des cas trop sauvages : on ne sait pas montrer que la correspondance obtenue préserve les facteurs ϵ de paires. C'est alors que Harris a eu l'idée (en écrivant [Ha2]) qu'on obtiendrait les propriétés voulues si l'on établissait des résultats semblables, mais en ne se limitant ni aux représentations cuspidales ni à la strate supersingulière. Cela demandait de revenir aux variétés du type de celles que nous regardons ici, parce que celles dont on vient de dire quelques mots admettent en w une fibre spéciale qui ne possède que la strate supersingulière, et sont associées à un groupe dont la composante en w est anisotrope modulo le centre (ses représentations correspondent donc seulement au spectre discret de $GL_n(F_w)$).

7.3. Pour expliquer l'idée, supposons démontré le théorème 2 de 3.6. On associe donc à chaque $\pi \in \mathcal{A}_n^0(K)$ une représentation galoisienne ℓ–adique de dimension n notée $\sigma_n(\pi)$; comme on l'a déjà dit ci-dessus, c'est une *bijection* de $\mathcal{A}_n^0(K)$ sur $\mathcal{G}_n^0(K)$ dont il faut voir qu'elle préserve les facteurs locaux de paires.

Étendons σ_n à des représentations non cuspidales de la façon suivante (ainsi que le suggère ce qui a été dit en 2.4) : soit $\pi \in \mathcal{A}_n(K)$ que l'on suppose être une induite irréductible de cuspidales : $\pi = \pi_1 \times \pi_2 \times \cdots \times \pi_t$, associée à une partition $n = n_1 + n_2 + \cdots + n_t$ (on laisse donc de côté les représentations obtenues comme des sous-quotients plus compliqués). Associons alors à une telle π la représentation suivante du groupe de Weil :

$$\tilde{\sigma}_n(\pi) = \sigma_{n_1}(\pi_1) \oplus \sigma_{n_2}(\pi_2) \oplus \cdots \oplus \sigma_{n_t}(\pi_t),$$

ainsi que $\tilde{r}_n(\pi) = \tilde{\sigma}_n(\pi^\vee) \otimes |\ |^{\frac{1-n}{2}}$.

PROPOSITION ([Ha2]).— *Considérons une représentation automorphe* Π *comme en* 7.1, *telle que sa composante locale* Π_w *soit une induite irréductible de cuspidales :* $\Pi_w = \pi = \pi_1 \times \pi_2 \times \cdots \times \pi_t$ *comme ci–dessus.*

Supposons que l'on sache prouver que, pour chaque Π *ainsi considérée, la semi-simplifiée de la restriction de* ρ_Π *au groupe de Weil local en* w *est isomorphe à* $\tilde{r}_n(\pi)$. *Alors les bijections* σ_n *préservent les facteurs* ϵ *de paires.*

Donnons quelques indications sur la preuve de ce résultat. Puisque les σ_n sont bijectives, on peut poser pour $\rho \in \mathcal{G}_n^0(K)$ et $\rho' \in \mathcal{G}_{n'}^0(K)$:

$$\epsilon_A(\rho \times \rho', s, \psi) = \epsilon(\sigma_n^{-1}(\rho) \times \sigma_{n'}^{-1}(\rho'), s, \psi).$$

Il s'agit de prouver l'identité :

$$\epsilon_A(\rho \times \rho', s, \psi) = \epsilon(\rho \otimes \rho', s, \psi).$$

Or on peut évidemment étendre (de façon unique) ϵ_A aux couples (ρ, ρ') de représentations semi-simples mais non nécessairement irréductibles de telle sorte que le résultat soit bi-additif en le couple (transforme somme directe sur l'un ou l'autre des facteurs en produit), une propriété déjà satisfaite par ϵ. Si l'on veut, on peut même étendre ϵ et ϵ_A aux représentations virtuelles.

En vertu du théorème d'induction de Brauer, il suffit alors d'établir l'égalité entre ϵ et ϵ_A pour deux représentations induites de caractères $\rho = \mathrm{Ind}_L^K(\chi)$ et $\rho' = \mathrm{Ind}_{L'}^K(\chi')$. Pour ce faire, on globalise la situation en considérant deux extensions $\tilde{L}/F$ et $\tilde{L}'/F$ (changeant éventuellement le corps global F !) se localisant à une place au-desssus de w en nos extensions locales, deux caractères $\tilde{\chi}$ et $\tilde{\chi}'$ des groupes des classes d'idèles admettant χ et χ' comme composantes en w, et tels que les conditions nécessaires soient remplies (voir 3.5) pour que les induites automorphes $\mathrm{IA}_{\tilde{L}}^F(\tilde{\chi})$ et $\mathrm{IA}_{\tilde{L}'}^F(\tilde{\chi}')$ existent et se descendent en des représentations automorphes Π et Π' sur $G(\mathbf{A})$, auxquelles sont associées des représentations ℓ–adiques $\tilde{\rho}$ et $\tilde{\rho}'$. On peut voir que dans ce cas les composantes locales Π_w (resp. Π'_w) sont des induites de cuspidales. Les représentations galoisiennes $\tilde{\rho}$ et $\tilde{\rho}'$ coïncident à torsion et passage aux contragrédientes près aux induites galoisiennes $\mathrm{Ind}_{\tilde{L}}^F(\tilde{\chi})$ et $\mathrm{Ind}_{\tilde{L}'}^F(\tilde{\chi}')$ et donc leurs facteurs locaux (de paires) coïncident ainsi (après torsion et passage aux contragrédientes) presque partout avec ceux de Π et Π' ; en vertu d'un argument déjà expliqué en 4.4, ces facteurs coïncident donc *partout*, et en particulier en w.

D'autre part, notre hypothèse appliquée à Π (resp. Π') nous dit que $\rho_w = \tilde{r}_n(\Pi_w)$ et que ${\rho'}_w = \tilde{r}_{n'}({\Pi'}_w)$. L'égalité des facteurs ϵ galoisiens et automorphes à la place w signifie donc que l'on a l'égalité voulue entre ϵ_A et ϵ pour un tel couple (ρ, ρ') de représentations induites.

Le théorème 1 résultera alors du th. 2 et du th. 3 ci–dessous. Pour $\mathcal{L}_\xi$ un faisceau ℓ–adique fixé comme en 4.3 sur nos variétés de Shimura, la limite inductive des groupes de cohomologie $H^i(X_{U^w,m} \times \overline{F}\ ,\ \mathcal{L}_\xi)$ est munie d'une action admissible du groupe $G(\mathbf{A}^\infty)$. Une représentation (irréductible admissible) Π^∞ de $G(\mathbf{A}^\infty)$ étant donnée,

on désigne par $R^i(\Pi^\infty)$ la composante isotypique correspondante et par $[R(\Pi^\infty)]$ la représentation virtuelle $(-1)^{n-1}\sum_i(-1)^i[R^i(\Pi^\infty)]$. En fait, cet usage des représentations virtuelles est commode parce que l'on veut appliquer des formules de traces. Toutefois, lorsque le plus haut poids de la représentation ξ est suffisamment régulier, on sait que la cohomologie est concentrée en degré $n-1$, et dans ce cas $[R(\Pi^\infty)]$ est simplement la classe de $R^{n-1}(\Pi^\infty)$; cette dernière représentation est non nulle si et seulement si Π^∞ est la composante finie d'une représentation automorphe d'un certain type archimédien (dépendant de ξ), et c'est alors un multiple par l'entier a (7.1) de la représentation ρ_Π associée.

Le théorème qui suit assure que les hypothèses de la proposition ci-dessus sont bien satisfaites. Il en résultera finalement qu'on a bien attrapé la correspondance de Langlands locale.

THÉORÈME 3.— *Soit Π^∞ une représentation admissible irréductible dont la composante locale en w est une induite de cuspidales $\Pi_w = \pi = \pi_1 \times \pi_2 \times \cdots \times \pi_t$ comme ci-dessus. Alors la représentation (virtuelle) $[R(\Pi^\infty)]$, restreinte au groupe de Weil local en w (ce qu'on note par un indice w), est donnée par la formule :*

$$n[R(\Pi^\infty)_w] = \dim([R(\Pi^\infty)])[\tilde{r}_n(\Pi_w)].$$

Ce dernier résultat proviendra lui-même du suivant. On ne suppose plus que la composante locale soit nécessairement une pure induite de cuspidales (i.e. on l'autorise à être un sous-quotient d'une induite réductible).

THÉORÈME 4.— *On a, dans le groupe de Grothendieck des représentations admissibles du produit $GL_n(F_w) \times W_{F_w}$, la formule :*

$$n[\Pi_w \otimes R(\Pi^\infty)_w] = \dim([R(\Pi^\infty)]) \sum_{h=0}^{n-1} \mathrm{Ind}_{P_h}^{GL_n(F_w)} \left\{ \sum_{\tau,\beta} m^{(h)}(\Pi_w;\tau,\beta)[(\Psi^*_{n-h}(\tau) \otimes |\ |^{-\frac{h}{2}}) \times \beta] \right\},$$

où l'on a utilisé les notations suivantes :

P_h désigne le sous-groupe parabolique de $GL_n(F_w)$ constitué des matrices triangulaires inférieures par blocs, avec deux blocs diagonaux de taille respective (en descendant vers la droite) $(n-h)\times(n-h)$ et $h\times h$ (cas particulier : $P_0 = GL_n(F_w)$) ; on note Ind *l'induction unitaire (d'une représentation du sous-groupe de Levi $GL_{n-h}(F_w) \times GL_h(F_w)$, étendue à P_h de façon usuelle).*

β décrit l'ensemble des classes de représentations admissibles irréductibles du groupe $GL_h(F_w)$ et τ celles du groupe multiplicatif du corps gauche D^*_{n-h} de centre F_w et d'invariant $1/(n-h)$. On a noté Ψ^*_{n-h} la représentation locale définie en 5.3 *(relative au corps local $K = F_w$) et $\Psi^*_{n-h}(\tau)$ la composante τ-isotypique* (5.5) *(c'est une représentation virtuelle du produit $GL_{n-h}(F_w) \times W_{F_w}$) ; la torsion $|\ |^{-\frac{h}{2}}$ porte sur le groupe de Weil.*

Enfin, $m^{(h)}(\Pi_w; \tau, \beta)$ désigne la multiplicité avec laquelle la représentation $\mathrm{JL}(\tau)^\vee \otimes \beta$ de $GL_{n-h}(F_w) \times GL_h(F_w)$ figure dans le module de Jacquet $\mathrm{J}_{U_h}(\Pi_w)$ (avec U_h le radical unipotent de P_h), augmenté de termes (qui n'interviennent pas dans l'application que nous avons en vue) correspondant à des composants $\alpha \otimes \beta$ de $\mathrm{J}_{U_h}(\Pi_w)$ tels que α soit non tempérée.

Rappelons que le module de Jacquet est l'espace des coïnvariants sous U_h , tordu par la racine carrée du module de P_h. Il est facile de déduire le th. 3 du th. 4. On utilise pour cela une formule tirée de [B-Z] donnant $\mathrm{J}_{U_h}(\pi)$ pour π une induite $\pi = \pi_1 \times \pi_2 \times \cdots \times \pi_t$ de cuspidales (associée à une partition $n = n_1 + n_2 + \cdots + n_t$) :

$$\mathrm{J}_{U_h}(\pi) = \sum_I (\times_{i \in I} \pi_i) \otimes (\times_{i \notin I} \pi_i) \ ,$$

(égalité dans le groupe de Grothendieck des représentations admissibles de longueur finie de $GL_{n-h}(F_w) \times GL_h(F_w)$) ; la somme est indexée par les sous-ensembles $I \subset \{1, 2, \cdots, n\}$ tels que $\sum_{i \in I} n_i = n - h$, de telle sorte que $(\times_{i \in I} \pi_i)$ (resp. $(\times_{i \notin I} \pi_i)$) est une représentation alors bien définie de $GL_{n-h}(F_w)$ (resp. $GL_h(F_w)$)), en fait irréductible (et indépendante de l'ordre des facteurs) ainsi qu'on le voit en utilisant la propriété analogue pour $\times_{1 \leq i \leq n} \pi_i$. On voit aussi que $(\times_{i \in I} \pi_i)$ ne peut appartenir à la série discrète que si I est réduit à un élément. Il en résulte que, sous l'hypothèse du théorème 3, ne peuvent intervenir dans la formule donnée par le th. 4 que des triples $\alpha = \pi_i$ (avec $n_i = n - h$) , $\tau = \mathrm{JL}(\alpha)^\vee$ et β tels que l'induite de $\alpha \times \beta$ redonne Π_w. Si l'on admet le th. 2 , on voit que la contribution galoisienne de ce triple est donnée par :

$$r_{n_i}(\pi_i)|\ |^{-\frac{n-n_i}{2}} = \sigma_{n_i}(\pi_i^\vee)|\ |^{\frac{1-n}{2}} \ ;$$

la somme de ces contributions est $\tilde{\sigma}_n(\pi^\vee)|\ |^{\frac{1-n}{2}} = \tilde{r}_n(\pi)$. Le théorème 3 résulte ainsi du théorème 4.

D'autre part le théorème 2 lui-même résulte aussi du théorème 4 : pour Π_w *cuspidale*, tous les modules de Jacquet non triviaux s'annulent et il ne reste plus dans la formule ci-dessus que le terme correspondant à $h = 0$, soit :

$$n[\Pi_w \otimes R(\Pi^\infty)_w] = \dim([R(\Pi^\infty)])(\Psi_n^*(\tau)) \ ,$$

avec $\tau = \mathrm{JL}(\Pi_w)^\vee$.

8. QUELQUES INDICATIONS SUR LA PREUVE DU THÉORÈME 4

8.1. Il s'agit de calculer la cohomologie ℓ-adique (relativement au faisceau $\mathcal{L}_\xi$) des variétés $X_{U^w,m} \times_F \overline{F}_w$. Pour cela on commence par étudier la géométrie de leur (mauvaise) réduction $\overline{X}_{U^w,m}$ et on calcule les faisceaux de cycles évanescents $\Phi^j \otimes \mathcal{L}_\xi$ en fonction des espaces locaux $\Psi^j_{(n-h)}$ du paragraphe 5. Nous noterons $\mathcal{H}^*_{\xi,U^w,m}$ la représentation virtuelle (dans un groupe de Grothendieck convenable) somme alternée (normalisée de sorte que H^{n-1} soit affecté du signe +) des $H^i(X_{U^w,m} \times_F \overline{F}_w \ , \ \mathcal{L}_\xi)$ et $\mathcal{H}^*_\xi$ la somme alternée des limites inductives. En utilisant la suite spectrale des cycles évanescents (7.1), on peut voir que cette représentation est une somme suivant h (variant de 0 à $n-1$) de sommes correspondant aux diverses strates $\overline{X}^{(h)}_{U^w,m}$:

$$\mathcal{H}^*_\xi = \sum_h \sum_{i,j} (-1)^{n-1+i+j} \varinjlim_{U^w,m} H^i_c(\overline{X}^{(h)}_{U^w,m} \times \overline{k} \ , \ \Phi^j \otimes \mathcal{L}_\xi),$$

où $\overline{k}$ désigne une clôture algébrique du corps résiduel k de $K = F_w$. Ceci 'explique' déjà la sommation sur h dans l'énoncé du th.4.

On peut aussi facilement donner une interprétation de l'induction depuis le parabolique P_h : en un point géométrique x de $\overline{X}^{(h)}_{U^w,m}$, la base de Drinfeld définit (avec les notations du paragraphe 6) un homomorphime surjectif de $(\varpi^{-m}\mathcal{O}/\mathcal{O})^n$ vers $(\mathcal{G}_A)_{\varpi^m}(x)$, lequel est isomorphe à $(\varpi^{-m}\mathcal{O}/\mathcal{O})^h$, d'où un noyau qui est un sous-$(\mathcal{O}/\varpi^m\mathcal{O})$- module libre de rang $(n-h)$ dans $(\varpi^{-m}\mathcal{O}/\mathcal{O})^n$. Ceci définit une partition de $\overline{X}^{(h)}_{U^w,m}$ en sous-variétés $\overline{X}^{(h)}_{U^w,m,M_m}$ (lieu où ce noyau coïncide avec un sous-module M_m donné.)

Prenons pour M_m le sous-module constitué des éléments de $(\varpi^{-m}\mathcal{O}/\mathcal{O})^n$ dont les h premières coordonnées sont nulles. Les sous-variétés correspondantes constituent un système projectif sur lequel opère le groupe P_h (ainsi que le groupe $G(\mathbf{A}^\infty)$ privé de sa composante en w). Posons alors :

$$\mathcal{H}^{i,j,(h)}_{\xi,M} = \varinjlim_{U^w,m} H^i_c(\overline{X}^{(h)}_{U^w,m,M_m} \times \overline{k} \ , \ \Phi^j \otimes \mathcal{L}_\xi) \ .$$

C'est un $G(\mathbf{A}^{w,\infty}) \times P_h \times W_K$ –module admissible, et il est assez formel de vérifier qu'il redonne par induction la cohomologie à support de $\overline{X}^{(h)}_{U^w,m}$ à valeurs dans $\Phi^j \otimes \mathcal{L}_\xi$: un élément c de ce dernier espace de cohomologie étant donné, considérer la fonction sur $GL_n(F_w)$ à valeurs dans $\mathcal{H}^{i,j,(h)}_{\xi,M}$ qui à g associe la restriction de gc au système projectif des $\overline{X}^{(h)}_{U^w,m,M_m}$ ...

On obtient de cette façon la formule :

$$\mathcal{H}^*_\xi = \sum_h \sum_{i,j} (-1)^{n-1+i+j} \mathrm{I}^{GL_n(K)}_{P_h} \mathcal{H}^{i,j,(h)}_{\xi,M}$$

où l'on a noté I l'induction naïve (non unitaire).

8.2. On étudie la fibre spéciale $\overline{X}_{U^w,m}$ au moyen des *variétés d'Igusa de première espèce* (généralisant les courbes du même nom qui servent dans la théorie de la réduction des courbes modulaires) ; on nomme ainsi, pour U^w , m et h comme ci-dessus, le revêtement étale galoisien (de groupe $GL_h(\mathcal{O}/\varpi^m\mathcal{O})$) :

$$I^{(h)}_{U^w,m} \to \overline{X}^{(h)}_{U^w,0}$$

qui classifie les isomorphismes (structures de niveau partielles) :

$$\alpha^{\mathrm{et}} : (\varpi^{-m}\mathcal{O}/\mathcal{O})^h \to (\mathcal{G}_A)^{\mathrm{et}}_{\varpi^m}$$

vers le quotient étale de $(\mathcal{G}_A)_{\varpi^m}$. Ce sont des variétés sur k. On les envoie dans la fibre spéciale $\overline{X}_{U^w,m}$ par une application j_* semilinéaire définie comme suit : à un point de $I^{(h)}_{U^w,m}$ à valeurs dans un k–schéma S, défini par des données $(A, \lambda, \iota, \bar{\eta}^p, \bar{\phi}_2, \cdots \bar{\phi}_r)$ comme au paragraphe 6, plus la donnée supplémentaire α^{et}, on associe le point de $\overline{X}_{U^w,0}$ suivant : $(A^{p^{mf(n-h)}}, \lambda^{p^{mf(n-h)}}, \iota, \bar{\eta}^p, \bar{\phi}_2, \cdots \bar{\phi}_r)$, (où f désigne le degré de k sur $\mathbf{F}_p$) ; puis on le munit de l'unique base de Drinfeld qui relève le composé de $\alpha^{p^{mf(n-h)}}$ avec la projection de $(\varpi^{-m}\mathcal{O}/\mathcal{O})^n$ sur $(\varpi^{-m}\mathcal{O}/\mathcal{O})^h$ donnée par les h dernières coordonnées ; un tel relèvement existe parce que l'extension $(\mathcal{G}_A)_{\varpi^m}$ de $(\mathcal{G}_A)^{\mathrm{et}}_{\varpi^m}$ par $(\mathcal{G}_A)^0_{\varpi^m}$ se scinde après le changement de base que nous avons effectué. Ces morphismes j_* vont en fait dans la sous-variété $\overline{X}_{U^w,m,M_m}$ et sont définis de telle façon

qu'ils s'insèrent dans des diagrammes commutatifs :

$$\begin{array}{ccc} I^{(h)}_{U^w,m} & \xrightarrow{j_*} & \overline{X}_{U^w,m,M_m} \\ \downarrow & & \downarrow \\ \overline{X}^{(h)}_{U^w,0} & \xrightarrow{\mathrm{Fr}^{mf(n-h)}} & \overline{X}^{(h)}_{U^w,0}. \end{array}$$

j_* ainsi défini est un isomorphisme sur le schéma réduit $\overline{X}^{\mathrm{red}}_{U^w,m,M_m}$.

Lorsque U^w et m varient, les $I^{(h)}_{U^w,m}$ constituent de façon évidente un système projectif (à morphismes de transition étales). Harris et Taylor considèrent parfois une variante de ce système projectif, avec le morphisme de $I^{(h)}_{U^w,m}$ vers $I^{(h)}_{U^w,m'}$ composé par $\mathrm{Fr}^{f(n-h)(m-m')}$, de façon à les rendre compatibles avec les morphismes j_*.

8.3. Les variétés d'Igusa admettent des *extensions formelles* $(I^{(h)}_{U^w,m})^\wedge$: on désigne ainsi l'extension étale (relevant $I^{(h)}_{U^w,m}/\overline{X}^{(h)}_{U^w,0}$) du schéma formel $(\overline{X}^{(h)}_{U^w,0})^\wedge$, restriction à $\overline{X}^{(h)}_{U^w,0}$ du complété de $X_{U^w,0}$ le long de sa fibre spéciale ; c'est aussi comme plus haut le classifiant – mais cette fois-ci au-desssus de $(\overline{X}^{(h)}_{U^w,0})^\wedge$ – des structures de niveau partielles (vers la partie étale). On dénote alors par $(I^{(h)}_{U^w,m})^\wedge(t)$ le schéma au-dessus de $(I^{(h)}_{U^w,m})^\wedge$ (un revêtement cette fois-ci très ramifié) qui classifie les bases de Drinfeld partielles :

$$\alpha^0 : (\varpi^{-m}\mathcal{O}/\mathcal{O})^{n-h} \to (\mathcal{G}_A)^0_{\varpi^t}$$

vers la partie formelle $(\mathcal{G}_A)^0_{\varpi^t}$ de $(\mathcal{G}_A)_{\varpi^t}$. La fibre spéciale réduite de $(I^{(h)}_{U^w,m})^\wedge(t)$ s'identifie à $I^{(h)}_{U^w,m}$.

On définit ensuite un morphisme $j^\wedge_* : (I^{(h)}_{U^w,m})^\wedge(m) \to (\overline{X}^{(h)}_{U^w,m,M_m})^\wedge$ étendant le morphisme j_* précédent. Pour cela on part de données $(A, \lambda, \iota, \bar{\eta}^p, \bar{\phi}_2, \cdots \bar{\phi}_r)$ comme ci-dessus (mais cette fois-ci sur un $(\mathcal{O}/\varpi^s\mathcal{O})$–schéma S), plus les données supplémentaires α^0 et α^{et} , et on doit leur associer une déformation de la structure déjà définie par j_*, avec en plus d'une base de Drinfeld : on prend comme déformation celle associée à la déformation $\mathcal{G}_A/(\mathcal{G}_A)^0_{\varpi^m}$ de $(\mathcal{G}_A)^{p^{mf(n-h)}}_{|\mathrm{Spec}(k)}$, munie de la base de Drinfeld

définie à partir de α^0, α^{et}, et du scindage :

$$(\mathcal{G}_A)^0_{\varpi^m} \times (\mathcal{G}_A)^{\mathrm{et}}_{\varpi^m} \to \left(\mathcal{G}_A/(\mathcal{G}_A)^0_{\varpi^m}\right)_{\varpi^m}$$

défini par :

$$(x, y) \to \varpi^{-m} x + y \ .$$

Le morphisme $j_*^\wedge$ ainsi obtenu est un *isomorphisme.* Noter que sa construction dépend du choix de l'uniformisante ϖ.

8.4. Il s'agit ensuite – et c'est là le point le plus fastidieux du manuscrit – de décrire des actions sur le système projectif des variétés d'Igusa. On a une action plus ou moins évidente du groupe $G(\mathbf{A}^{p,\infty})$; rappelons d'autre part que $G(\mathbf{Q}_p)$ est isomorphe au produit $\mathbf{Q}_p^* \times GL_n(F_w) \times \prod_{i\geq 2} B_{w_i}^*$.

Harris et Taylor définissent une action de $\mathbf{Q}_p^* \times \left(GL_{n-h}(F_w) \times GL_h(F_w)\right)^+ \times \prod_{i\geq 2} B_{w_i}^*$ sur le système projectif des variétés d'Igusa $(I^{(h)}_{U^w,m})^\wedge(m)$, où $\left(GL_{n-h}(F_w) \times GL_h(F_w)\right)^+$ désigne le sous-monoïde du produit constitué des couples (g, g') pour lesquels il existe $a \in F_w^*$ tel que $a^{-1}g^{-1}$ et ag' soient tous deux à coefficients entiers. L'action de $\mathbf{Q}_p^*$ se factorise en fait via la valuation. D'autre part, l'action sur le système projectif des fibres spéciales réduites $I^{(h)}_{U^w,m}$ se factorise (via $\mathrm{val.det}(g)$) à travers $\mathbf{Q}_p^* \times \left(\mathbf{Z} \times GL_h(F_w)\right)^+ \times \prod_{i\geq 2} B_{w_i}^*$, où $\left(\mathbf{Z} \times GL_h(F_w)\right)^+$ désigne le sous-monoïde défini de manière analogue, comme l'ensemble des couples (c, g') tels que $\varpi^E g'$ soit à coefficients entiers, avec E la partie entière de $-c/(n-h)$.

Notons $\tilde{\mathrm{Fr}}_q$ l'élément $(q^{-1}, -1, 1)$ de $\mathbf{Q}_p^* \times \left(\mathbf{Z} \times GL_h(F_w)\right)^+$. On vérifie que l'action sur les $I^{(h)}_{U^w,m}$ de cet élément coïncide avec celui du Frobenius $\mathrm{Fr}_q = \mathrm{Fr}^f$. Ceci est une généralisation au cas présent de la classique 'relation de congruence', sous la forme que lui a donnée Piatetski-Shapiro [PS]. $\tilde{\mathrm{Fr}}_q^{n-h}$ se relève en l'élément $(q^{h-n}, \varpi^{-1}, 1)$ de $\mathbf{Q}_p^* \times \left(GL_{n-h}(F_w) \times GL_h(F_w)\right)^+$, noté $\tilde{\mathbf{Fr}}_q^{n-h}$ (qui dépend du choix de ϖ) ; la notation est trompeuse car $\tilde{\mathbf{Fr}}_q$ n'existe pas.

La relation entre les morphismes $j_*^\wedge$ et les actions d'un élément

$$\gamma \in G(\mathbf{A}^{p,\infty}) \times \mathbf{Q}_p^* \times \left(GL_{n-h}(F_w) \times GL_h(F_w)\right)^+ \times \prod_{i\geq 2} B_{w_i}^* \ \subset \ G(\mathbf{A}^\infty)$$

sur chacun des systèmes projectifs des $(I^{(h)}_{U^w,m})^\wedge(m)$ et $(\overline{X}^{(h)}_{U^w,m,M_m})^\wedge$ se traduit par la commutativité des diagrammes :

$$\begin{array}{ccc} (I^{(h)}_{U^w,m})^\wedge(m) & \xrightarrow{\gamma\ \tilde{\mathbf{Fr}}_q^{(n-h)(m-m')}} & (I^{(h)}_{U'^w,m'})^\wedge(m') \\ \downarrow j_*^\wedge & & \downarrow j_*^\wedge \\ (\overline{X}^{(h)}_{U^w,m,M_m})^\wedge & \xrightarrow{\gamma} & (\overline{X}^{(h)}_{U'^w,m',M_{m'}})^\wedge \end{array}$$

(chaque fois que U^w, m, U'^w, m' sont tels que cela ait un sens). Noter que pour $\gamma = 1$ on retrouve après restriction aux fibres spéciales la commutativité du diagramme écrit plus haut.

8.5. Considérons les faisceaux de cycles évanescents sur $I^{(h)}_{U^w,m}$ que la théorie de Berkovich associe au schéma formel $(I^{(h)}_{U^w,m})^\wedge(t)$. Nous les notons encore $\Phi^j(t)$. Le théorème de comparaison de Berkovich permet de voir qu'ils coïncident (lorsque $t = m$), via les morphismes $j_*^\wedge$, avec les faisceaux du même nom définis (par la théorie usuelle) sur $\overline{X}^{(h)}_{U^w,m,M_m}$. Les actions que nous (n') avons (pas) définies ci-dessus permettent de faire opérer

$$G(\mathbf{A}^{p,\infty}) \times \mathbf{Q}_p^* \times \big(GL_{n-h}(F_w) \times GL_h(F_w)\big)^+ \times \prod_{i\geq 2} B_{w_i}^*$$

sur la limite inductive

$$H^i_c\big(I^{(h)}\ , \Phi^j \otimes \mathcal{L}_\xi\big) = \varinjlim_{U_w,m,t} H^i_c\big(I^{(h)}_{U^w,m} \otimes \overline{k}\ ,\ \Phi^j(t) \otimes \mathcal{L}_\xi\ \big);$$

et c'est une action admissible. Comme on peut voir que l'action de $\tilde{\mathbf{Fr}}_q^{(n-h)}$ est inversible, et que l'adjonction de cet élément permet d'engendrer le groupe $G(\mathbf{A}^{p,\infty}) \times \mathbf{Q}_p^* \times GL_{n-h}(F_w) \times GL_h(F_w) \times \prod_{i\geq 2} B_{w_i}^*$ tout entier, notre action se prolonge à ce dernier. Enfin, ces espaces se trouvent également munis d'une action du groupe de Weil W_{F_w}.

On obtient des isomorphismes :

$$H^i_c\big(I^{(h)}\ , \Phi^j \otimes \mathcal{L}_\xi\big) \simeq \mathcal{H}^{i,j,(h)}_{\xi\ ,M}$$

par image réciproque par les applications $\tilde{\mathbf{Fr}}_q^{(n-h)m} \circ (j_*^\wedge)^{-1}$ de $(I^{(h)}_{U^w,m})^\wedge(m)$ vers $(\overline{X}^{(h)}_{U^w,m,M_m})^\wedge$. Cette torsion par $\tilde{\mathbf{Fr}}_q^{(n-h)m}$ possède la vertu, d'après ce qu'on a vu plus haut, de rendre ces isomorphismes compatibles avec l'action du groupe considéré ci-dessus. On montre enfin que l'action sur $H^i_c\big(I^{(h)}\ , \Phi^j \otimes \mathcal{L}_\xi\big)$ du radical unipotent de P_h est triviale.

De tout cela, il résulte que l'on peut récrire la formule donnée en 8.1 sous la forme suivante :

$$\mathcal{H}^*_\xi = \sum_h \sum_{i,j} (-1)^{n-1+i+j} \mathrm{I}^{GL_n(K)}_{P_h} H^i_c\big(I^{(h)}\ , \Phi^j \otimes \mathcal{L}_\xi\big)\ .$$

8.6. L'objectif est maintenant de calculer les faisceaux $\Phi^j(t)$ en fonction des espaces locaux $\Psi^j_{(n-h),t}$ du paragraphe 5. Si x est un point géométrique algébrique de $I^{(h)}_{U^w,m}$, alors le complété de l'anneau local en x s'identifie à l'espace des déformations $\mathrm{Def}_t(\mathcal{G}_x)$ du groupe p–divisible $\mathcal{G}_x = (\mathcal{G}_A)_x$, munies d'une base de Drinfeld partielle α^0 de niveau t (l'autre α^{et} se déforme uniquement). Cet espace de déformations est fibré sur l'espace des déformations (avec base de Drinfeld) $\mathrm{Def}_t(\mathcal{G}^0_x)$ de la partie formelle $\mathcal{G}^0_x$. Vu comme schéma au-dessus de $\mathrm{Def}_t(\mathcal{G}^0_x)$, notre espace $\mathrm{Def}_t(\mathcal{G}_x)$ classifie les extensions du groupe constant $\mathcal{G}^{\mathrm{et}}$ par la déformation universelle $\tilde{\mathcal{G}}^0$ et un calcul classique montre que cet espace d'extensions est isomorphe à $\big(\tilde{\mathcal{G}}^0\big)^h$; en particulier $\mathrm{Def}_t(\mathcal{G}_x)$ est formellement lisse au-dessus de $\mathrm{Def}_t(\mathcal{G}^0_x)$; on en déduit que le calcul de $\Phi^j(t)$ se ramène au calcul des cycles évanescents pour $\mathrm{Def}_t(\mathcal{G}^0_x)$, lequel est isomorphe au schéma formel $\mathrm{Spf}(R_{K,(n-h),t})$ du paragraphe 5. Tout choix d'un isomorphisme entre $\mathcal{G}^0_x$ et $\Sigma_{(n-h)}$ induit un isomorphisme entre $\mathrm{Def}_t(\mathcal{G}^0_x)$ et $\mathrm{Spf}(R_{K,(n-h),t})$, d'où un isomorphisme entre $\Phi^j(t)$ et $\Psi^j_{(n-h),t}$.

Parce que le groupe $\mathcal{G}^0$ n'est pas constant sur les variétés d'Igusa $I^{(h)}_{U^w,m}$ (ni même sur leur limite), un tel isomorphisme n'existe pas globalement. Pour remédier à cela, on introduit des revêtements visant à trivialiser partiellement $\mathcal{G}^0$. Les *variétés d'Igusa de deuxième espèce* sont des revêtements étales galoisiens :

$$J^{(h)}_{U^w,m,s} \longrightarrow I^{(h)}_{U^w,m} \otimes_k \overline{k}$$

de groupe $\big(\mathcal{O}_{D_{(n-h)}}/\varpi^s \mathcal{O}_{D_{(n-h)}}\big)^*$.

Ils sont définis en considérant pour chaque s le classifiant Isom_s, au-dessus de $I^{(h)}_{U^w,m} \otimes_k \overline{k}$, des isomorphismes entre le groupe $(\Sigma_{(n-h)})_{\varpi^s}$ des points de ϖ^s–torsion

du groupe $\Sigma_{(n-h)}$ (vu par changement de base comme groupe sur $I^{(h)}_{U^w,m} \otimes_k \overline{k}$) et le groupe $(\mathcal{G}^0)_{\varpi^s}$. Pour $s' \geq s$, on dispose de morphismes évidents de $\mathrm{Isom}_{s'}$ dans Isom_s, dont on note Y_s l'intersection des images (s' variant). Finalement, $J^{(h)}_{U^w,m,s}$ est défini comme le schéma réduit Y_s^{red} associé. D'une façon un peu vague, il classifie les trivialisations 'modulo ϖ^s' de $\mathcal{G}^0$.

Lorsque s varie, les $J^{(h)}_{U^w,m,s}$ constituent un système projectif de revêtements étales de $I^{(h)}_{U^w,m} \otimes_k \overline{k}$, dont le groupe de Galois est $\mathcal{O}^*_{D_{(n-h)}}$. Cela permet d'associer (de manière usuelle) à chaque représentation ρ (ℓ–adique admissible) de ce groupe un $\overline{\mathbf{Q}}_\ell$–faisceau lisse $\mathcal{F}(\rho)$, obtenu comme quotient du faisceau constant de fibre ρ sur la 'limite projective' (mais cela se formule aisément sur un cran fini) par l'action de $\mathcal{O}^*_{D_{(n-h)}}$. En particulier, si on se souvient que $\Psi^j_{(n-h),t}$ est muni d'une action de $\mathcal{O}^*_{D_{(n-h)}}$, on peut former le faisceau associé $\mathcal{F}(\Psi^j_{(n-h),t})$. Explicitons ce qui se passe en un point géométrique algébrique x de $I^{(h)}_{U^m,m} \otimes_k \overline{k}$: on considère tous les relèvements possibles x_∞ de x à la limite projective (c'est-à-dire les systèmes compatibles de relèvements aux crans finis) ; ils sont permutés par $\mathcal{O}^*_{D_{(n-h)}}$. Par définition, la donnée d'un élément de la fibre $\left(\mathcal{F}(\Psi^j_{(n-h),t})\right)_x$ correspond à la donnée pour chaque relèvement x_∞ d'un élément de $\Psi^j_{(n-h),t}$, de façon compatible à l'action de $\mathcal{O}^*_{D_{(n-h)}}$.

Mais par construction des $J^{(h)}_{U^w,m,s}$, la donnée de x_∞ correspond à la donnée d'un isomorphisme entre $\mathcal{G}^0_x$ et $\Sigma_{(n-h)}$; d'après ce qu'on a dit plus haut, cela nous donne un isomorphisme entre $(\Phi^j(t))_x$ et $\Psi^j_{(n-h),t}$. Un élément de $(\Phi^j(t))_x$ étant fixé, on associe ainsi à chaque x_∞ un élément de $\Psi^j_{(n-h),t}$, de façon compatible à l'action de groupe (par définition de l'action de $\mathcal{O}^*_{D_{(n-h)}}$ sur $\Psi^j_{(n-h),t}$). On voit finalement que l'on a obtenu de la sorte un isomorphisme canonique entre les fibres :

$$(\Phi^j(t))_x \simeq \left(\mathcal{F}(\Psi^j_{(n-h),t})\right)_x .$$

Berkovich a fourni une preuve du fait que cet isomorphisme de fibres provient d'un isomorphisme entre les faisceaux $\Phi^j(t)$ et $\mathcal{F}(\Psi^j_{(n-h),t})$. Il s'agit d'un passage à la limite assez délicat qui utilise sa théorie (laquelle ne s'applique pas directement car les limites projectives que nous considérons ne sont pas de type fini).

Décomposons maintenant l'action de $\mathcal{O}^*_{D_{(n-h)}}$ sur $\Psi^j_{(n-h),t}$ en la somme des composantes isotypiques $\Psi^j_{(n-h),t}(\rho)$ (pour t fixé, seul un nombre fini interviennent), ρ

décrivant l'ensemble des classes d'équivalence de représentations admissibles irréductibles de $\mathcal{O}^*_{D_{(n-h)}}$ (ou, ce qui revient au même, de représentations de $D^*_{(n-h)}$ modulo torsion non ramifiée). Il en résulte une décomposition du faisceau associé :

$$\mathcal{F}(\Psi^j_{(n-h),t}) = \bigoplus_\rho \mathcal{F}(\rho) \otimes \Psi^j_{(n-h),t}(\rho) \ ,$$

le second facteur du produit tensoriel étant vu comme un faisceau constant. Joint à l'isomorphisme ci-dessus, cela permet d'exprimer autrement les espaces de cohomologie des faisceaux $\Phi^j \otimes \mathcal{L}_\xi$ sur les variétés d'Igusa :

$$H^i_c(I^{(h)} \ , \ \Phi^j \otimes \mathcal{L}_\xi) = \bigoplus_\rho \ H^i_c\big(I^{(h)} \ , \ \mathcal{F}(\rho) \otimes \mathcal{L}_\xi\big) \otimes \Psi^j_{(n-h),t}(\rho),$$

où l'on a noté $H^i_c\big(I^{(h)} \ , \ \mathcal{F}(\rho) \otimes \mathcal{L}_\xi\big)$ la limite inductive des $H^i_c\big(I^{(h)}_{U^w,m} \otimes \overline{k} \ , \ \mathcal{F}(\rho) \otimes \mathcal{L}_\xi\big)$. Ceci permet de donner une autre forme à la formule de (8.1) exprimant $\mathcal{H}^*_\xi$:

$$\mathcal{H}^*_\xi = \sum_\rho \sum_h \sum_{i,j} (-1)^{n-1+i+j} \ \mathrm{I}^{GL_n(K)}_{P_h} \{H^i_c\big(I^{(h)} \ , \mathcal{F}(\rho) \otimes \mathcal{L}_\xi\big) \otimes \Psi^j_{(n-h)}(\rho)\} \ .$$

Soit encore :

$$\mathcal{H}^*_\xi = \sum_\rho \sum_h \mathrm{I}^{GL_n(K)}_{P_h} \mathcal{H}^{(h)}_{\xi,\rho} \otimes \Psi^*_{(n-h)}(\rho) \ ,$$

avec $\mathcal{H}^{(h)}_{\xi,\rho}$ la somme alternée $\sum(-1)^{h+i}(H^i_c\big(I^{(h)} \ , \ \mathcal{F}(\rho) \otimes \mathcal{L}_\xi\big)$.

8.7. Actions

Harris et Taylor relèvent l'action du semi-groupe

$$G(\mathbf{A}^{p,\infty}) \times \mathbf{Q}^*_p \times \big(\mathbf{Z} \times GL_h(F_w)\big)^+ \times \prod_{i \geq 2} B^*_{w_i}$$

sur le système projectif des variétés d'Igusa $I^{(h)}_{U^w,m}$ en une action de

$$G(\mathbf{A}^{p,\infty}) \times \mathbf{Q}^*_p \times \big(\mathbf{Z} \times GL_h(F_w)\big)^+ \times \prod_{i \geq 2} B^*_{w_i} \times D^*_{(n-h)}$$

sur le système des $J^{(h)}_{U^w,m,s}$. Il en résulte que les faisceaux $\mathcal{F}(\rho)$ sont équivariants sous l'action du semi–groupe ci–dessus, qui opère donc sur les groupes de cohomologie $\mathcal{H}^{(h)}_{\xi,\rho}$.

Comme plus haut, cette action se prolonge au groupe $G(\mathbf{A}^{p,\infty}) \times \mathbf{Q}_p^* \times \mathbf{Z} \times GL_h(F_w) \times \prod_{i\geq 2} B_{w_i}^*$. Rappelons d'autre part que $\Psi_{(n-h)}^*(\rho)$ est un $(GL_{(n-h)}(F_w) \times W_{F_w})$ – module virtuel.

Munissons alors $\mathcal{H}_{\xi,\rho}^{(h)} \otimes \Psi_{(n-h)}^*(\rho)$ d'une action de

$$G(\mathbf{A}^{p,\infty}) \times \mathbf{Q}_p^* \times GL_{(n-h)}(F_w) \times GL_h(F_w) \times \prod_{i\geq 2} B_{w_i}^* \times W_{F_w}$$

ainsi définie : un élément $(g^p, k, g_w^0, g_w^{\mathrm{et}}, g_{w_i}, \tau)$ agit sur le premier facteur du produit tensoriel par $(g^p, k|\tau|, \mathrm{val.det}(g_w^0) - \mathrm{val.}(\tau), g_w^{\mathrm{et}}, g_{w_i})$ et par (g_w^0, τ) sur le second ; on a noté $\mathrm{val.}(\tau)$ et $|\tau|$ la valuation et la valeur absolue de l'élément de K^* qui correspond à τ par l'isomorphisme de la théorie du corps de classes local. Par induction, on obtient donc une action (virtuelle) de

$$G(\mathbf{A}^{\infty}) \times W_{F_w} = G(\mathbf{A}^{p,\infty}) \times \mathbf{Q}_p^* \times \times GL_n(F_w) \times \prod_{i\geq 2} B_{w_i}^* \times W_{F_w}$$

qui correspond, par l'égalité exprimée en 8.6, à celle que l'on a sur $\mathcal{H}_{\xi}^*$.

8.8. La preuve du théorème 4 est ainsi ramenée au calcul de la représentation $\mathcal{H}_{\xi,\rho}^{(h)}$ du groupe $G(\mathbf{A}^{p,\infty}) \times \mathbf{Q}_p^* \times \mathbf{Z} \times GL_h(F_w) \times \prod_{i\geq 2} B_{w_i}^*$: il s'agit de la limite inductive de cohomologies à support de variétés lisses (mais non propres) sur $\overline{k}$, à valeurs dans un $\overline{\mathbf{Q}}_\ell$–faisceau lisse. Si ϕ est une fonction sur l'algèbre de Hecke du groupe ci-dessus (privé de son facteur $\mathbf{Z}$), on peut la translater par une puissance assez grande de l'élément $\tilde{\mathrm{Fr}}_q$ considéré plus haut ; notons $\tilde{\mathrm{Fr}}_q^N(\phi)$ la fonction qui à g associe $\phi(\tilde{\mathrm{Fr}}_q^{-N} g)$. Harris et Taylor obtiennent une description à la Langlands–Kottwitz (cf. [Ko1], [Ko2]) de l'ensemble des points sur $\overline{k}$ des variétés $J_{U^w,m,s}^{(h)}$, avec les différentes actions mentionnées ci-dessus. La formule des traces de Lefschetz permet alors le calcul de la trace de l'action de $\tilde{\mathrm{Fr}}_q^N(\phi)$ pour N assez grand. On utilise ici le fait, conjecturé par Deligne et prouvé par Fujiwara [Fu], qu'il existe une formule des traces simple et sans terme à l'infini (pour la cohomologie à support), valide après torsion par une assez grande puissance de Frobenius.

On utilise ensuite la formule des traces de Selberg pour calculer la trace sur l'espace des formes automorphes d'une fonction $\phi'(N)$ sur le groupe

$$G(\mathbf{A}) = G(\mathbf{A}^{p,\infty}) \times \mathbf{Q}_p^* \times GL_n(F_w) \times \prod_{i\geq 2} B_{w_i}^* \times G(\mathbf{R})$$

obtenue ainsi : supposant la fonction ϕ considérée ci-dessus décomposée en le produit d'une fonction $\phi^{w,\infty}$ sur $G(\mathbf{A}^{p,\infty}) \times \prod_{i\geq 2} B^*_{w_i}$, d'une fonction ϕ_0 sur $\mathbf{Q}_p^*$ et enfin d'une fonction ϕ_w^{et} sur $GL_h(F_w)$, on pose :

$$\phi'(N) = \phi^{w,\infty} \times \phi'_w(\phi_w^{\text{et}}, \rho, -N) \times \phi_0^{(N)} \times \phi_{\infty,\xi}$$

où $\phi_0^{(N)}(x) = \phi_0(q^N x)$, où $\phi_{\infty,\xi}$ est l'habituelle fonction sur $G(\mathbf{R})$ qui isole les représentations figurant dans la cohomologie de $\mathcal{L}_\xi$, tandis que $\phi'_w(\phi_w^{\text{et}}, \rho, -N)$ est une fonction sur $GL_n(F_w)$ (dépendant de ϕ_w^{et} , de ρ et de N) qui possède la vertu suivante : pour π une représentation irréductible de $GL_n(F_w)$ et π' une représentation irréductible de $GL_{(n-h)}(F_w)$, notons $\pi^{(h)}[\pi']$ la représentation de $GL_h(F_w)$ telle que $\pi' \otimes \pi^{(h)}[\pi']$ soit la composante π'–isotypique du module de Jacquet $J_{U_h}(\pi)$ semi-simplifié. On a alors :

$$\text{tr}(\pi)(\phi'_w(\phi_w^{\text{et}}, \rho, -N)) = \sum_\mu \mu(-N)\text{tr}(\pi^{(h)}[\text{JL}^\vee(\rho) \otimes (\mu \circ \text{val.det})])(\phi_w^{\text{et}}) \ ,$$

où μ décrit l'ensemble des quasi–caractères de $\mathbf{Z}$.

La comparaison des deux formules donne pour N grand l'égalité, à un facteur près, entre la trace (donnée par la formule de Selberg) de $\phi'(N)$ sur l'espace des formes automorphes et celle (donnée par Lefschetz) de $\tilde{\text{Fr}}_q^N(\phi)$ sur la cohomologie $\mathcal{H}^{(h)}_{\xi,\rho}$. Elle utilise divers ingrédients : des résultats de [D-K-V] sur la correspondance de Jacquet–Langlands, des résultats de Labesse ([L1], [L2]), en particulier sur le fait que la trace de certaines fonctions sépare différents sous–quotients d'une induite, de Clozel ([Cl1], [Cl2]) sur les représentations automorphes des groupes unitaires tordus, de Clozel-Labesse ([C-L]) ...

8.9. L'égalité de traces ainsi obtenue se traduit alors par la proposition qui suit.

Fixons une représentation admissible $\Pi^{w,\infty}$ du groupe $G(\mathbf{A}^{p,\infty}) \times \prod_{i\geq 2} B^*_{w_i}$. Notons $\mathcal{H}^{(h)}_{\xi,\rho}(\Pi^{w,\infty})$ la composante isotypique correspondante, qui est donc une représentation du groupe $\mathbf{Q}_p^* \times \mathbf{Z} \times GL_h(F_w)$. D'autre part, notons $\mathcal{H}^*_\xi(\Pi^{w,\infty})$ la composante isotypique dans la représentation globale $\mathcal{H}^*_\xi$: il s'agit cette fois d'une représentation de $\mathbf{Q}_p^* \times GL_n(F_w)$. Décomposons le module de Jacquet (relativement à P_h) de cette dernière (toujours après semi-simplification) :

$$J_{U_h}(\mathcal{H}^*_\xi(\Pi^{w,\infty})) = \sum_{\chi,\alpha,\beta} m^{(h)}_\xi(\chi,\alpha,\beta)\chi \otimes \alpha \otimes \beta \ ,$$

avec χ (resp. α, resp. β) décrivant les représentations admissibles de $\mathbf{Q}_p^*$ (resp. $GL_{(n-h)}(F_w)$, resp. $GL_h(F_w)$), et $m_\xi^{(h)}(\chi,\alpha,\beta)$ la multiplicité de $\chi\otimes\alpha\otimes\beta$.

PROPOSITION.— *On a la formule :*

$$\mathcal{H}_{\xi,\rho}^{(h)}(\Pi^{w,\infty}) = n^{-1}\sum_{\chi,\alpha,\beta,\mu} m_\xi^{(h)}(\chi,\alpha,\beta)(\mathrm{tr}\alpha)(\phi_{\mathrm{JL}^\vee(\rho\otimes(\mu^{-1}\circ\mathrm{oval.det})}[\chi\otimes\mu\otimes\beta],$$

où χ,α,β sont comme ci-dessus, et où μ décrit l'ensemble des quasi-caractères de $\mathbf{Z}$; d'autre part $\phi_{\mathrm{JL}^\vee(\rho)}$ est un pseudo-coefficient de la représentation $\mathrm{JL}^\vee(\rho)$: pour α tempérée, $(\mathrm{tr}\alpha)(\phi_{\mathrm{JL}^\vee(\rho)})$ vaut 1 si α est isomorphe à $\mathrm{JL}^\vee(\rho)$ et 0 sinon.

On combine alors cette expression de $\mathcal{H}_{\xi,\rho}^{(h)}(\Pi^{w,\infty})$ avec le résultat de 8.6, qui donne $\mathcal{H}_\xi(\Pi^{w,\infty})$ comme la somme des induites des $\mathcal{H}_{\xi,\rho}^{(h)}(\Pi^{w,\infty})\otimes\Psi_{(n-h)}^*(\rho)$, pour obtenir finalement le théorème 4.

BIBLIOGRAPHIE

[A-C] J. Arthur et L. Clozel - *Simple algebras, base change and the advanced theory of the trace formula*, Annals of Math. Studies **120**, Princeton University Press (1989).

[Ba] A. Badulescu - *Correspondance de Jacquet-Langlands pour les corps locaux de caractéristique non nulle*, thèse, Université de Paris-Sud (janvier 1999).

[Be1] V.G. Berkovich - *Etale cohomology for non-archimedean analytic spaces*, Publ. Math. I.H.E.S. **78** (1993), 5–161.

[Be2] V.G. Berkovich - *Vanishing cycles for formal schemes*, Invent. Math. **115** (1994), 539–571.

[Be3] V.G. Berkovich - *Vanishing cycles for formal schemes II*, Invent. Math. **125** (1996), 367–390.

[B-H] C.J. Bushnell et G. Henniart - *Davenport–Hasse relations and an explicit Langlands correspondence, II : an interim report*, prépublication (Janvier 99).

[Bo] J.-F. Boutot - *Uniformisation p–adique des variétés de Shimura*, Séminaire Bourbaki exp. 831 (1996–1997), Astérisque **245** (1997), 307–322.

[Boy] P. Boyer - *Mauvaise réduction des variétés de Drinfeld et conjecture de Langlands locale*, thèse, Université de Paris 11 Orsay (Oct. 1998).

[Br] J. Brylinski, *Un lemme sur les cycles évanescents en dimension relative 1*, Appendice à [Ca2].

[B-Z] I.N. Bernstein et A.V. Zelevinsky - *Induced representations of reductive $\mathfrak{p}$-adic groups I*, Ann. Sci. ENS **10** (1977), 441–472.

[Ca1] H. Carayol - *Sur la mauvaise réduction des courbes de Shimura*, Comp. Math. **59** (1986), 151–230.

[Ca2] H. Carayol - *Sur les représentations ℓ-adiques associées aux formes modulaires de Hilbert*, Ann. Sci. ENS **19** (1986), 409–467.

[Ca3] H. Carayol - *Non-abelian Lubin-Tate theory*, dans 'Automorphic forms, Shimura varieties, and L-functions' vol II (L. Clozel et J.S. Milne, eds), Academic Press (1990).

[Ca4] H. Carayol - *Variétés de Drinfeld compactes, d'après Laumon, Rapoport et Stuhler*, Séminaire Bourbaki exp. 756 (1991–1992), Astérisque **206** (1992), 369–409.

[Cart] P. Cartier - *La conjecture de Langlands pour $GL(2)$ et la démonstration de Ph. Kutzko*, Séminaire Bourbaki exp. 550 (1979–1980), Springer Lecture Notes in Math. **842** (1981), 112–138.

[Cl1] L. Clozel - *Représentations galoisiennes associées aux représentations automorphes autoduales de $GL(n)$*, Publ. Math. IHES **73** (1991), 97–145.

[Cl2] L. Clozel - *On the cohomology of Kottwitz's arithmetic varieties*, Duke Math. J. **72** (1993), 757–795.

[Cl3] L. Clozel - *Nombre de points des variétés de Shimura sur un corps fini (d'après R. Kottwitz)*, Séminaire Bourbaki, exp. 766 (1992–1993), Astérisque **216** (1993), 121–149.

[C-L] L. Clozel et J.-P. Labesse - *Changement de base pour les représentations cohomologiques de certains groupes unitaires*, Appendice à [L2], prépublication.

[De1] P. Deligne - *Travaux de Shimura*, Séminaire Bourbaki exp. 389 (1970–1971), Springer Lecture Notes in Math. **244** (1971), 123–165.

[De2] P. Deligne - *Les constantes des équations fonctionnelles des fonctions L*, dans : Modular functions of one variable II, Springer Lectures Notes in Math. **349** (1973).

[De3] P. Deligne - Lettre à Piatetski-Shapiro (1973).

[D-K-V] P. Deligne, D. Kazhdan et M.-F. Vignéras - *Représentations des algèbres simples p-adiques*, dans I.-N. Bernstein, P. Deligne, D. Kazhdan et M.-F. Vignéras : 'Représentations des groupes réductifs sur un corps local', Hermann (1984).

[Dr1] V.G. Drinfeld - *Elliptic modules*, Math. Sbornik **23** (1974), 561–592.

[Dr2] V.G. Drinfeld - *Coverings of p-adic symmetric domains*, Funct. Anal. Appl. **10** (1976), 107–115.

[Fa] G. Faltings - *The trace formula and Drinfeld's upper halfplane*, Duke Math. J.

76 (1994), 467–481.

[Fu] K. Fujiwara - *Rigid geometry, Lefschetz–Verdier trace formula and Deligne's conjecture*, Invent. Math. **127** (1997), 489–533.

[G-J] P. Godement et H. Jacquet - *Zeta functions of simple algebras*, Springer Lectures Notes in Math. **260** (1972).

[Ha1] M. Harris - *Supercuspidal representations in the cohomology of Drinfeld upper half spaces : elaboration of Carayol's program*, Invent. Math. **129** (1997), 75–120.

[Ha2] M. Harris - *The local Langlands conjecture for $GL(n)$ over a p–adic field, $n < p$*, Invent. Math. **134** (1998), 177–210.

[H-T] M. Harris et R. Taylor - *On the geometry and cohomology of some simple Shimura varieties*, version préliminaire (1998).

[He1] G. Henniart - *La conjecture de Langlands locale pour $GL(3)$*, Mémoire S.M.F. n° 11/12 (1984).

[He2] G. Henniart - *La conjecture de Langlands locale numérique pour $GL(n)$*, Ann. Sci. ENS **21** (1988), 497–544.

[He3] G. Henniart - *Caractérisation de la correspondance de Langlands locale par les facteurs ϵ de paires*, Invent. Math. **113** (1993), 339–350.

[He4] G. Henniart - *Une preuve simple des conjectures de Langlands pour GL_n sur un corps p–adique*, manuscrit (Jan. 1999).

[H-H] G. Henniart et R. Herb - *Automorphic induction for GL_n (over local nonarchimedean fields)*, Duke Math. J. **78** (1995), 131–192.

[JPSS1] H. Jacquet, I.I. Piatetski-Shapiro, J. Shalika - *Facteurs L et ϵ du groupe linéaire*, C.R.A.S. Paris t. 289 (1980), 59–61.

[JPSS2] H. Jacquet, I.I. Piatetski-Shapiro, J. Shalika - *Rankin–Selberg convolutions*, Am. J. of Math. **105** (1983), 367–483.

[K-M] N.M. Katz et B. Mazur - *Arithmetic moduli of elliptic curves*, Annals of Math. Studies **108**, Princeton University Press (1985).

[Ko1] R. Kottwitz - *Points on some Shimura varieties over finite fields*, J. of the A.M.S. **5** (1992), 373–444.

[Ko2] R. Kottwitz - *On the λ–adic representations associated to some simple Shimura varieties*, Invent. Math. **108** (1992), 653–665.

[Ku] Ph. Kutzko - *The Langlands conjecture for $GL(2)$ of a local field*, Annals of Math. **112** (1980), 381–412.

[L1] J.P. Labesse - *Fonctions élémentaires et lemme fondamental pour le changement de base stable*, Duke Math. J. **61** (1990), 519–530.

[L2] J.P. Labesse - *Changement de base conditionnel pour les groupes réductifs*, prépublication.

[La] R.P. Langlands - *On the Zeta-functions of some simple Shimura varieties* Canadian J. of Math **31** n° 6 (1979), 1121–1216.

[L-R-S] G. Laumon, M. Rapoport et U. Stuhler - *$\mathcal{D}$-elliptic sheaves and the Langlands correspondence*, Invent. Math. **113** (1993), 217–338.

[PS] I.I. Piatetski-Shapiro - *Zeta functions of modular curves*, dans : Modular functions of one variable II, Springer Lectures Notes in Math. **349** (1973).

[R] F. Rodier - *Représentations de $GL(n,k)$ où k est un corps p-adique*, Séminaire Bourbaki exp. 587 (1981–1982), Astérisque **92–93** (1982), 201–218.

[Ro] J. Rogawski, *Representations of $GL(n)$ and division algebras over a p-adic field*, Duke Math. J. **50** (1983), 161–196.

[R-Z] M. Rapoport et T. Zink, *Period spaces for p-divisible groups*, Annals of Math. Studies **141**, Princeton University Press (1996).

[S-S] P. Schneider and U. Stuhler, *The cohomology of p-adic symmetric spaces*, Invent. Math. **105** (1991), 47–122.

Henri CARAYOL

Institut de Recherche Mathématique Avancée
Université Louis Pasteur et C. N. R. S.
7, rue René Descartes
F-67084 Strasbourg Cedex
adresse électronique : carayol@math.u-strasbg.fr

Astérisque

OLIVIER MATHIEU

Classification des algèbres de Lie simples

Astérisque, tome 266 (2000), Séminaire Bourbaki,
exp. n° 858, p. 245-286

<http://www.numdam.org/item?id=SB_1998-1999__41__245_0>

© Société mathématique de France, 2000, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (http://smf4.emath.fr/Publications/Asterisque/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Séminaire BOURBAKI
51ème année, 1998-99, n° 858

Mars 1999

CLASSIFICATION DES ALGÈBRES DE LIE SIMPLES

par **Olivier MATHIEU**

Sommaire :

0. Introduction

1. Généralités sur les algèbres de Lie restreintes
   1.1. Identité de Jacobson
   1.2. Dérivations
   1.3. Exemples et remarques
2. Algèbres de Lie simples classiques
   2.1. Définition des algèbres de Lie simples classiques
   2.2. Un exemple de trop bonne réduction
   2.3. La $\mathbf{Z}$-forme canonique de $\mathfrak{g}_{\mathbf{C}}(I)$
3. Algèbres de Cartan restreintes
   3.1. Définition des algèbres de Cartan restreintes
   3.2. Un critère de simplicité
   3.3. L'algèbre de Lie $W(n, \mathbf{1})$
   3.4. La cohomologie de De Rham de X
   3.5. Un critère de "restreignabilité"
   3.6. L'algèbre de Lie $S(n, \mathbf{1})$, pour $n \geq 3$
   3.7. Algèbres de Poisson
   3.8. L'algèbre de Lie $H(2m, \mathbf{1})$
   3.9. Un autre critère de simplicité
   3.10. L'algèbre de Lie $K(2m+1, \mathbf{1})$
4. Le théorème de Block et Wilson
5. Défaut au lemme de Darboux et une classe d'algèbres de Lie non restreintes
   5.1. Notations pour la section 5
   5.2. Lemme de Darboux

5.3. Remarques sur les algèbres filtrées
5.4. Formes volumes sur X
5.5. Formes symplectiques
5.6. Exemples d'algèbres hamiltoniennes non restreintes
6. Algèbres de Cartan de niveau arbitraire
6.1. Algèbres à puissances divisées
6.2. Algèbres de Cartan de dimension infinie
6.3. Définition des algèbres de Cartan
7. La classification des algèbres de Lie simples
8. Quelques techniques de preuve
8.1. Sous-algèbres de Cartan
8.2. Construction de grosses sous-algèbres
8.3. Éléments sandwich
8.4. Théorèmes de reconnaissance
Bibliographie

0. INTRODUCTION

Au début du siècle, Killing et Cartan ont classifié les algèbres de Lie simples de dimension finie sur $\mathbf{C}$ ou, plus généralement, sur un corps algébriquement clos de caractéristique zéro. La classification fait apparaître quatre familles infinies $\mathfrak{sl}(l+1)$, $\mathfrak{so}(2l+1)$, $\mathfrak{sp}(2l)$, $\mathfrak{so}(2l)$, et cinq autres algèbres de Lie appelées G_2, F_4, E_6, E_7 et E_8. Pour les besoins de cet exposé, toutes ces algèbres de Lie seront appelées classiques, bien que la terminologie usuelle la réserve pour les quatre séries infinies. Il n'y a pas de familles continues d'algèbres de Lie simples ou, ce qui est équivalent, pour chaque entier n, il n'existe qu'au plus un nombre fini d'algèbres de Lie simples de dimension n. Afin d'éviter les répétitions, *nous sous-entendrons que toute algèbre de Lie est de dimension finie*, sauf indication explicite du contraire.

Le problème de la classification des algèbres de Lie simples de dimension finie sur un corps algébriquement clos K de caractéristique p a résisté plus longtemps. À partir des années 30, au départ sous l'impulsion de Jacobson, il est apparu qu'en plus des réductions mod. p des algèbres de Lie classiques (appelées encore algèbres de Lie classiques), il existe de nombreuses algèbres de Lie simples de dimension finie, analogues aux pseudo-groupes considérés par Cartan [**Cn**]. Ces algèbres de Lie sont appelées algèbres de Cartan. Dans la terminologie moderne, les pseudo-groupes sont certaines algèbres de Lie de dimension infinie, et les algèbres de Cartan sont donc des analogues de dimension finie en caractéristique finie de structures de dimension

infinie en caractéristique zéro. On trouvera dans [**Cn**],[**Kc1**], [**G**], [**Ma**] des énoncés analogues aux résultats exposés ici et qui concernent des classifications d'algèbres de Lie simples de dimension infinie en caractéristique zéro.

Bien que la question de la classification des algèbres de Lie simples ait été un problème clairement posé dès les années 30, ce n'est que dans les années 60 que toutes les classes d'algèbres de Cartan ont été finalement construites, et Kostrikin et Shafarevitch ont pu formuler la conjecture que toute algèbre de Lie simple est classique ou de Cartan [**KS**]. En raison de la structure de l'identité de Jacobi, il semble difficile d'espérer une classification en caractéristique 2 et 3, et en fait, pour ces caractéristiques très petites, il est facile d'obtenir des algèbres de Lie simples qui ne soient pas de Cartan ou classiques. Il est plus surprenant qu'en caractéristique 5, il existe une famille étrange d'algèbres de Lie simples, appelées algèbres de Melikian [**Me**], qui ne sont ni classiques ni de Cartan. En revanche, il n'existe pas d'autres exemples en caractéristique $p \geq 7$.

Une famille importante d'algèbres de Lie est celle des algèbres de Lie restreintes : ce sont les algèbres de Lie L telles que pour tout $x \in L$, $ad(x)^p$ soit une dérivation intérieure. La première étape de la classification a été obtenue par R. Block et R. Wilson en 1988, qui ont classifié les algèbres de Lie simples restreintes en caractéristique $p > 7$ [**BW1**]. Comme en caractéristique zéro, il n'y a pas de famille continue d'algèbres de Lie simples restreintes.

La classification des algèbres de Lie simples non nécessairement restreintes en caractéristique > 7 a été le point d'orgue d'un article de Strade et de Wilson ([**SW**]) et elle est basée sur de nombreux papiers, en particulier ceux de Strade [**St1-7**]. Les experts espèrent que ce résultat peut s'étendre tel quel en caractéristique 7 et, en ajoutant la famille de Melikian, en caractéristique 5. La classification des algèbres de Lie générales est plus complexe que dans le cas des algèbres de Lie restreintes, car elle contient des familles continues d'algèbres de Lie simples. Elle est aussi moins explicite. Nous reviendrons sur ce point plus loin.

Le but principal de cet exposé est de décrire cette classification. Dans un souci de clarté, nous avons cherché à utiliser systématiquement le formalisme du calcul différentiel pour simplifier de nombreux calculs explicites trouvés dans la littérature.

L'organisation de cet exposé écrit est la suivante. La section 1 est consacrée à la théorie des algèbres de Lie restreintes, notion cruciale par la suite. La section 2 porte sur la définition des algèbres classiques : puisque l'idée de base est claire, nous avons surtout insisté sur des points techniques, et cette section n'a pas vraiment un caractère introductif (on devrait pouvoir la sauter en première lecture). Dans la

section 3, nous contruisons les algèbres de Cartan restreintes : dans cette section, nous nous sommes attaché à décrire en détail ces algèbres et nous avons donné des preuves complètes du calcul de leur série dérivée, et de la simplicité du dernier terme de la série dérivée. Il nous semble que l'utilisation de cohomologie de De Rham pour déterminer les séries dérivées fournit une approche plus simple que les calculs en coordonnées locales trouvés dans la littérature. Nous espérons ne pas avoir introduit d'erreurs dans cette section en voulant préciser ou simplifier certains énoncés. La section 4 porte sur le théorème de Block et Wilson. Initialement, la section 5 était surtout destinée à illustrer la complexité de la classification des algèbres de Lie restreintes. Puis nous nous sommes rendu compte qu'il était possible de décrire plus explicitement les algèbres de Cartan par le formalisme introduit dans cette section : faute de temps, il n'a pas été possible d'expliquer cette description dans le cadre le plus général, et une version plus complète fera l'objet d'une publication séparée. La section 6 expose la définition des algèbres de Cartan non restreintes, et elle s'inspire de [**W1**] et [**St7**]. La section 7 contient le résultat principal, aboutissement des travaux de Strade et Wilson. Dans la section 8, nous avons cherché à expliquer les idées les plus simples mises en œuvre dans la classification : en réalité le théorème de classification est très compliqué.

Remerciements : Nous remercions O. Gabber, V. G. Kac et J.-P. Serre pour d'utiles commentaires sur cet exposé.

Conventions : *Sauf indication contraire, K désigne un corps algébriquement clos de caractéristique finie p, et à partir de la section* 3 *on supposera* $p > 3$.

Notations : *Pour une algèbre de Lie L, nous noterons $Z(L)$ son centre et nous poserons* $L^{(1)} = [L, L]$, $L^{(2)} = [L^{(1)}, L^{(1)}]$ *etc.. et* $L^{(\infty)} = \cap_{k \geq 1} L^{(k)}$.

1. GÉNÉRALITÉS SUR LES ALGÈBRES DE LIE RESTREINTES

Dans cette section, nous allons rappeler la définition des algèbres de Lie restreintes. Je crois que la motivation initiale pour la notion d'algèbre de Lie restreinte était d'établir un analogue de la théorie de Galois pour des extensions de corps purement inséparables. Soit K un corps, et soit $Der(K)$ l'algèbre de Lie des dérivations de K. En fait Jacobson [**J1**] a établi une correspondance entre :

(i) les extensions $K \subset K'$ avec $K' \subset K^{1/p}$, et

(ii) les sous-K espaces vectoriels de $Der(K)$ qui sont des sous-algèbres de Lie restreintes.

La correspondance associe à K' l'algèbre de Lie des dérivations $\delta : K \to K$ telles que $\delta(f^p) = 0$ pour tout $f \in K'$. Les structures d'algèbres de Lie restreintes constituent aussi un point crucial dans la définition de l'opérateur de Cartier (voir exemple 8) et dans une conjecture de Grothendieck (voir exemple 9).

1.1. Identité de Jacobson

Il est connu que l'on a l'identité $(x + y)^p = x^p + y^p$ dans toute $\mathbf{F}_p$-algèbre commutative. Jacobson a trouvé une généralisation de cette identité valable pour toute $\mathbf{F}_p$-algèbre associative mais non nécessairement commutative. Commençons par rappeler la définition des polynômes de Lie, plus exactement les polynômes de Lie en deux variables. Fixons un anneau de base k (dans la suite, $k = \mathbf{Z}$, $k = \mathbf{Q}$ où k est un corps K de caractéristique p).

Un polynôme en deux variables commutatives x, y est un élément $P(x, y) \in k[x, y]$. Pour toute k-algèbre commutative A et tous $a, b \in A$, on peut évaluer l'expression $P(a, b) \in A$. Notons $k < X, Y >$ l'algèbre tensorielle en les deux variables X et Y (i.e. la k-algèbre associative libre en les deux indéterminées X et Y) et soit $Lib_k(X, Y)$ la k-algèbre de Lie libre en X et Y. Plus concrètement, $Lib_k(X, Y)$ est la sous-algèbre de Lie de $k < X, Y >$ engendrée par X et Y. Par définition, un *polynôme de Lie en deux variables* X, Y est un élément $P(X, Y)$ de $Lib_k(X, Y)$. Pour toute k-algèbre de Lie L et $A, B \in L$, on peut définir son évaluation $P(A, B) \in L$. Pour en donner une définition formelle, on remarque qu'il existe un unique morphisme d'algèbres de Lie $\phi : Lib_k(X, Y) \to L$ avec $\phi(X) = A$ et $\phi(Y) = B$ et on définit $P(A, B) = \phi(P(X, Y))$. Un exemple de polynôme de Lie est fourni par la formule de Campbell-Hausdorff :

$$\exp tX \exp tY = \exp\left(\sum_{n \geq 1} \frac{t^n}{n} H_n(X, Y)\right)$$

qui est valable dans l'algèbre $\mathbf{Q}[[t]] < X, Y >$. Dans cette identité, $H_n(X, Y)$ est un polynôme de Lie à coefficients entiers p-adiques pour tout $n \leq p$.

THÉORÈME 1 (Formule de Jacobson).— *Soit p un nombre premier. Il existe un (unique) polynôme de Lie $J_p(X, Y) \in Lib_{\mathbf{F_p}}(X, Y)$, tel que :*

$$(x + y)^p = x^p + y^p + J_p(x, y),$$

pour tous éléments x, y d'une $\mathbf{F_p}$-algèbre associative.

Preuve : Considérons le facteur de t^p dans la formule de Campbell-Hausdorff et multiplions-le par $p!$. On obtient : $(X + Y)^p = X^p + Y^p + (p - 1)!H_p(X, Y) + p\tau$, où

$\tau \in \mathbf{Z} < X, Y >$. Définissons $J_p(X,Y)$ comme la réduction modulo p de $-H_p(X,Y)$. Puisque $(p-1)! = -1 \bmod. p$, l'identité $(x+y)^p = x^p + y^p + J_p(x,y)$ est valide dans toute $\mathbf{F_p}$-algèbre associative. Enfin l'unicité de $J_p(X,Y)$ résulte du fait que l'application naturelle $Lib_{\mathbf{F_p}}(X,Y) \to \mathbf{F_p} < X, Y >$ est injective. C.Q.F.D.

Pour la preuve originelle de Jacobson, on pourra se reporter à **[J2]**.

Exemples : Pour $p = 2$, $(x+y)^2 = x^2 + y^2 + [x,y]$, d'où

$$J_2(X,Y) = [X,Y].$$

Pour $p = 3$, $[x,[x,y]] = x^2y - 2xyx + yx^2 = x^2y + xyx + yx^2$, d'où :

$$J_3(X,Y) = [X[X,Y]] + [Y[Y,X]].$$

Pour $p = 5$, on a :

$$J_5(X,Y) = [X[X[X[X,Y]]]] + 2[Y[X[X[Y,X]]]] + 2[X[Y[Y[X,Y]]]] + [Y[Y[Y[Y,X]]]].$$

1.2. Dérivations

Soit R une $\mathbf{F}_p$-algèbre non nécessairement commutative ou associative (par exemple une algèbre de Lie, une algèbre d'octonions). Une dérivation est une application linéaire $\delta : R \to R$ qui satisfait à l'identité de Leibniz : $\delta(x.y) = (\delta x).y + x.\delta\, y$. On a donc $\delta^n(x.y) = \sum_{a+b=n} \binom{n}{a} (\delta^a\, x).\delta^b\, y$, et en particulier $\delta^p(x.y) = (\delta^p x).y + x.\delta^p\, y$. Par conséquent, l'ensemble $Der\, R$ des dérivations de R est une algèbre de Lie avec une structure supplémentaire, à savoir l'application $\delta \mapsto \delta^p$: cette structure est appelée une *p-structure*, une notion dont la définition formelle suit. Une algèbre de Lie munie d'une p-structure est dite *restreinte*.

Soit L une algèbre de Lie sur un corps K de caractéristique p. Pour tout $x \in L$, $ad(x)^p$ est une dérivation. Nous dirons qu'une algèbre de Lie est "*restreignable*" si $ad(x)^p$ est une dérivation intérieure pour tout $x \in L$ (l'utilisation de cet horrible néologisme nous a semblé particulièrement commode pour cet exposé). Grâce à l'identité de Jacobson, on remarque qu'une algèbre de Lie est "restreignable" dès qu'elle possède une base $(x_i)_{i \in I}$ telle que $ad(x_i)^p$ soit intérieure pour tout $i \in I$. Cependant, il n'est pas suffisant de supposer que $ad(x_i)^p$ est intérieure pour un système $(x_i)_{i \in I}$ de générateurs de l'algèbre de Lie L. Une *p-structure* est une application $x \in L \mapsto x^{(p)} \in L$ telle que $ad(x^{(p)}) = ad(x)^p$ satisfaisant aux conditions de compatibilité suivantes :

(i) $(\lambda x)^{(p)} = \lambda^p\, x^{(p)}$

(ii) $(x+y)^{(p)} = x^{(p)} + y^{(p)} + J_p(x,y)$

pour tous $\lambda \in K$ et $x, y \in L$.

Toute algèbre de Lie "restreignable" L admet une p-structure. En effet soit $(x_i)_{i\in I}$ une base de L et soit $(y_i)_{i\in I} \in L$ avec $ad(x_i)^p = ad(y_i)$, alors il existe une unique p-structure $x \mapsto x^{(p)}$ telle que $x_i^{(p)} = y_i$. Si l'on choisit un ordre total sur I, la p-structure est donnée par $(\sum_{i\in I} \lambda_i x_i)^{(p)} = \sum_{i\in I} \lambda_i^p y_i + J(\lambda_i x_i, \sum_{j<i} \lambda_j x_j)$. Pour vérifier que cette application est bien une p-structure, il suffit de remarquer que J_p est un cocycle, i.e. satisfait à $J(x+y,z) + J(x,y) = J(y,z) + J(x,y+z)$. Comme il est clair que deux p-structures diffèrent d'une application p-linéaire de L dans $Z(L)$, une algèbre de Lie "restreignable" et de centre trivial admet une unique p-structure. C'est pourquoi nous dirons qu'une algèbre de Lie de centre trivial est *non restreinte* si elle n'est pas "restreignable".

1.3. Exemples et remarques

1) Si V est un espace vectoriel, l'algèbre de Lie $\mathfrak{gl}(V)$ des endomorphismes de V est restreinte. Plus généralement n'importe quelle algèbre associative A est une algèbre de Lie restreinte : la p-structure est $x \mapsto x^p$. Si V est de dimension finie, l'algèbre de Lie $\mathfrak{sl}(V)$ des endomorphismes de trace nulle est restreinte. En effet $Tr(x^p) = Tr(x)^p$.

2) Pour n'importe quelle algèbre R, l'agèbre de Lie $Der(R)$ est restreinte. Par exemple l'algèbre de Lie W_X de tous les champs de vecteurs d'une variété X est restreinte.

3) Soit X une variété et β une forme. L'algèbre de Lie $W_X(\beta)$ de tous les champs de vecteurs qui préservent β, c'est-à-dire les champs de vecteurs ξ tels que $\xi.\beta = 0$ est restreinte.

4) Soit G un groupe algébrique et $\mathfrak{g}$ son algèbre de Lie. Alors $\mathfrak{g}$ est restreinte. En effet, $\mathfrak{g}$ est l'espace des champs de vecteurs ξ sur G qui sont invariants relativement aux translations à gauche par G. Or si ξ est G-invariant à gauche, ξ^p est aussi G-invariant à gauche.

5) La définition de p-structure n'impose aucune condition de compatibilité avec le crochet de Lie. En fait pour toute p-structure, l'identité (due également à Jacobson) $[x^{(p)}, y^{(p)}] = -ad^{p-1}(x)ad^p(y)(x)$ est satisfaite. Pour toute algèbre de Lie $\mathfrak{g}$, il existe une algèbre de Lie restreinte $L \supset \mathfrak{g}$ telle que $\mathfrak{g}$ soit un idéal de L et $L/\mathfrak{g}$ soit commutative. Par exemple, il suffit de prendre pour L la plus petite sous-algèbre de Lie restreinte de $U(\mathfrak{g})$ engendrée par $\mathfrak{g}$. Si $(x_i)_{i\in I}$ est une base de $\mathfrak{g}$, $(x_i^{p^n})_{i\in I, n\geq 0}$ est une base de L.

6) Soit L une algèbre de Lie, et soit Z son centre. Alors si L est restreinte,

L/Z est restreinte, car la p-structure passe au quotient. Réciproquement si L/Z est restreinte, L est "restreignable".

En revanche la dérivée $L^{(1)} = [L, L]$ d'une algèbre de Lie restreinte n'est en général pas restreinte, ni "restreignable". Avant d'en voir des exemples concrets en sections 5 et 6, donnons-en un exemple abstrait. Soit $\mathfrak{g}$ une algèbre de Lie non "restreignable" avec $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ et soit $L \subset U(\mathfrak{g})$ l'algèbre de Lie construite dans l'exemple précédent. Il est clair que $[L, L] = \mathfrak{g}$, donc $L^{(1)}$ n'est pas restreinte.

7) Soit L une algèbre de Lie. L'ensemble L' des éléments $x \in L$ tels que $ad(x)^p$ soit intérieure est un sous-espace vectoriel invariant par tout automorphisme de L, mais ce n'est en général ni un idéal ni même une sous-algèbre de Lie. On peut trouver des exemples où L' engendre L bien que $L' \neq L$.

Néanmoins, si l'algèbre de Lie est engendrée (en tant qu'algèbre) par ses éléments x tels que $ad(x)^{\frac{p+1}{2}} = 0$, alors l'algèbre de Lie est restreinte. En effet pour tout $x \in L$ tel que $ad(x)^{\frac{p+1}{2}}$, l'exponentielle tronquée :

$$\exp ad(x) = \sum_{0 \leq i < \frac{p+1}{2}} ad(x)^i$$

est un automorphisme d'algèbres de Lie. On en déduit que l'algèbre de Lie est linéairement engendrée par les éléments x qui satisfont à $ad(x)^{\frac{p+1}{2}} = 0$, et donc L est "restreignable". Cette hypothèse est satisfaite pour toutes les algèbres de Lie classiques si $p > 5$, puisque avec les notations usuelles (voir section 2) on a $ad(e_i)^4 = ad(f_i)^4 = 0$.

8) Soit X une variété affine lisse, soit W_X l'algèbre de Lie des champs de vecteurs, soit Ω_X l'espace des formes différentielles et soit d la différentielle extérieure. Il est clair que l'existence de d est intimement liée à la structure d'une application bilinéaire $W_X \times W_X \to W_X$ (le crochet des champs de vecteurs) et la relation $d^2 = 0$ à l'identité de Jacobi. Puisque que W_X est aussi muni d'une application unaire $W_X \to W_X$ (i.e. la p-structure), on doit s'attendre à ce que Ω_X soit muni d'une structure supplémentaire. Cette structure supplémentaire a été inventée par Cartier [**Cr**] : c'est un isomorphisme d'anneaux $Car : \Omega_X^* \to H_{DR}^*(X)$, défini par $Car(f) = [f^p]$ et $C\,df = [f^{p-1}df]$ pour tout $f \in K[X]$ (ici $[\beta]$ désigne la classe de cohomologie de la forme fermée β).

9) On reprend les notations de l'exemple 8. Soit E un fibré vectoriel sur X, et ∇ une connexion. On sait que la première obstruction à trouver des sections horizontales à E est la non-nullité de la courbure R de la connexion : $R(\xi, \eta) = \nabla_{[\xi,\eta]} - [\nabla_\xi, \nabla_\eta]$ pour tous $\xi, \eta \in W_X$. La courbure traduisant la structure d'algèbre de Lie de W_X, on s'attend à une autre obstruction liée à la p-structure : c'est la p-courbure R_p définie par $R_p(\xi) = \nabla_{\xi^p} - \nabla_\xi^p$.

La nullité de R équivaut au fait que $\xi \mapsto \nabla_\xi$ est un morphisme d'algèbres de Lie. Lorsque $R = 0$, la p-courbure est additive en ξ et $[\nabla_\xi, R_p(\eta)] = 0$ pour tous $\xi, \eta \in W_X$. La nullité simultanée de R et de R_p équivaut au fait que $\xi \mapsto \nabla_\xi$ est un morphisme d'algèbres de Lie restreintes.

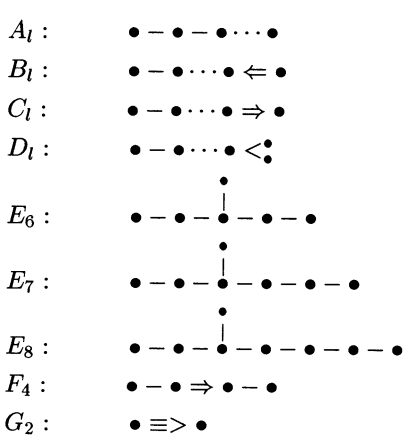
Signalons une conjecture fascinante de Grothendieck : Soit Y une variété affine lisse sur un corps de caractéristique 0 et soit E un fibré vectoriel sur X muni d'une connexion ∇ de courbure nulle. Si les réductions modulo p de ∇ sont des connexions de p-courbure nulle pour $p \gg 0$, la monodromie de ∇ est finie (voir par exemple **[Kz]**).

2. ALGÈBRES DE LIE SIMPLES CLASSIQUES

En première approximation, les algèbres de Lie classiques sont les réductions modulo p des algèbres de Lie simples définies sur $\mathbf{C}$. Dans ce paragraphe, nous allons préciser comment ces algèbres de Lie sont définies.

2.1. Définition des algèbres de Lie classiques

Soit $\mathfrak{g}$ une algèbre de Lie simple sur $\mathbf{C}$. Il existe un ouvert dense $\Omega \subset \mathfrak{g}$ tel que le noyau de $ad(x)$ ait une dimension constante l pour tout $x \in \Omega$. Ce nombre l est appelé le *rang* de $\mathfrak{g}$. On peut aussi le définir comme la dimension minimale des noyaux de $ad(x)$. Par un procédé que nous allons rappeler, on associe à $\mathfrak{g}$ un diagramme de Dynkin ayant l sommets, c'est-à-dire l'un des dessins suivants :



Notons I l'ensemble des sommets du diagramme de Dynkin. On définit la *matrice de Cartan* $A = (a_{i,j})_{i,j \in I}$ de I, par les conditions suivantes : $a_{i,i} = 2$ et pour $i \neq j$,

$a_{i,j} = a_{j,i} = 0$ si i et j ne sont pas connectés,

$a_{i,j} = a_{j,i} = -1$ si i et j sont connectés comme : $^i\bullet - \bullet^j$

$a_{i,j} = -1$, $a_{j,i} = -2$ si i et j sont connectés comme : $^i\bullet \Leftarrow \bullet^j$

$a_{i,j} = -1$, $a_{j,i} = -3$ si i et j sont connectés comme : $^i\bullet \equiv > \bullet^j$.

Définissons $\mathfrak{g}_{\mathbf{C}}(I)$ comme l'agèbre de Lie engendrée par les générateurs $(e_i, f_i, h_i)_{i\in I}$ et définie par les relations suivantes :

(i) $[h_i, h_j] = 0$,

(ii) $[h_i, e_j] = a_{i,j}e_j$,

(iii) $[h_i, f_j] = -a_{i,j}f_j$,

(iv) $[e_i, f_i] = h_i$,

(v) $[e_i, f_j] = 0$ pour $i \neq j$,

(vi) $ad(e_i)^{1-a_{i,j}}(e_j) = 0$ pour $i \neq j$,

(vii) $ad(f_i)^{1-a_{i,j}}(f_j) = 0$ pour $i \neq j$,

pour tous $1 \leq i, j \leq l$. Les deux dernières relations sont appelées *relations de Serre*. Le diagramme de Dynkin de $\mathfrak{g}$ est l'unique diagramme de Dynkin I tel que $\mathfrak{g} \simeq \mathfrak{g}_{\mathbf{C}}(I)$. Les algèbres simples de diagramme de Dynkin A_l ($l \geq 1$), B_l ($l \geq 2$), C_l ($l \geq 3$), D_l ($l \geq 4$), E_6, E_7, E_8, F_4 et G_2 constituent une famille complète de classes d'isomorphisme d'algèbres de Lie simples sur $\mathbf{C}$.

Rappelons que l'algèbre de type A_l est $\mathfrak{sl}(l+1)$, celle de type B_l est $\mathfrak{so}(2l+1)$, celle de type C_l est $\mathfrak{sp}(2l)$ et celle de type D_l est $\mathfrak{so}(2l)$. On impose $l \geq 2$ (respectivement $l \geq 3$, $l \geq 4$) pour le type B_l (respectivement pour le type C_l, D_l) pour éviter les répétitions (isomorphismes en basse dimension) ou pour éviter certaines algèbres qui ne sont pas simples. En effet les algèbres de type B_1, C_1 et A_1 sont isomorphes (i.e. $\mathfrak{so}(3) \simeq \mathfrak{sp}(2) \simeq \mathfrak{sl}(2)$), les algèbres de type C_2 et B_2 sont isomorphes (i.e. $\mathfrak{sp}(4) \simeq \mathfrak{so}(5)$), les algèbres de type D_1 et D_2 ne sont pas simples (car $\mathfrak{so}(2)$ est commutative et $\mathfrak{so}(4)$ est isomorphe à $\mathfrak{sl}(2) \times \mathfrak{sl}(2)$) et enfin les algèbres de type D_3 et A_4 sont isomorphes (i.e. $\mathfrak{so}(6) \simeq \mathfrak{sl}(4)$). Les experts de la caractéristique zéro appellent classiques les algèbres simples de type $ABCD$ et exceptionnelles les cinq autres. Dans la terminologie des experts de caractéristique finie, elles sont toutes appelées *algèbres de Lie classiques*, et nous suivrons ici cette terminologie.

On notera que cette présentation ne fait intervenir que des nombres entiers. Si nous considérons l'algèbre de Lie définie par cette présentation sur un corps de caractéristique finie, on trouve parfois des algèbres de trop grande dimension, et parfois de dimension infinie (ce problème ne se pose qu'en petite caractéristique).

C'est pourquoi la méthode pour obtenir les algèbres de Lie modulo p consiste à définir une $\mathbf{Z}$-forme de $\mathfrak{g}_{\mathbf{C}}(I)$. Rappelons qu'une $\mathbf{Z}$-*forme* d'une $\mathbf{C}$-algèbre quelconque

L est un sous-anneau Λ de L qui soit un $\mathbf{Z}$-module libre et tel que $\mathbf{C} \otimes \Lambda = L$. La $\mathbf{Z}$-forme de $\mathfrak{g}_{\mathbf{C}}(I)$ la plus simple à définir est le sous-anneau de Lie $\mathfrak{g}_{\mathbf{Z}}(I)$ engendré par les éléments e_i et f_i pour tout $i \in I$. Comme nous le verrons à la fin de cette section, cette forme est la forme canonique, à 6-torsion près. Pour tout corps K, on pose $\mathfrak{g}_K(I) = K \otimes \mathfrak{g}_{\mathbf{Z}}(I)$. Le fait suivant sera prouvé à la fin de cette section :

LEMME 2.— *Supposons $p > 3$. L'algèbre de Lie $\mathfrak{g}_K(I)$ est restreinte. De plus si I n'est pas de type A_l avec $l = -1$ modulo p, l'algèbre de Lie $\mathfrak{g}_K$ est simple.*

L'algèbre de Lie de type A_{np-1} est isomorphe à $\mathfrak{sl}(np)$. La matrice identique 1 est de trace nulle, et l'on pose $\mathfrak{psl}(np) = \mathfrak{sl}(np)/K1$. Il est facile de voir que $\mathfrak{psl}(np)$ est simple.

DÉFINITION 3.— *Soit K un corps de caractéristique > 3. Les algèbres de Lie classiques sont les algèbres $\mathfrak{g}_K(I)$ où I n'est pas de type A_l avec $l = -1$ modulo p et les algèbres $\mathfrak{psl}(np)$.*

2.2. Un exemple de trop bonne réduction

La définition de $\mathfrak{g}_K(I)$ peut sembler compliquée, puisque nous ne sommes intéressé qu'à la définition de $\mathfrak{g}_K(I)$ sur un corps algébriquement clos : en fait le processus de réduction en caractéristique finie fait partie de la définition. Voici un exemple d'une réduction modulo p de $\mathfrak{sl}(p)$ qui fournit une algèbre simple non classique.

Sur l'espace $E = \mathbf{F}_p^2$, choisissons une forme symplectique non nulle à valeur dans $\mathbf{F}_p$ notée $\alpha, \beta \in E \mapsto < \alpha | \beta >$. Soit $q \in \mathbf{C}$ une racine p-ième de l'unité. Notons $\mathcal{A}$ la $\mathbf{C}$-algèbre associative de base $(e_\alpha)_{\alpha \in E}$ dont le produit est donné par $e_\alpha e_\beta = q^{<\alpha|\beta>} e_{\alpha+\beta}$. Il est facile de voir que $\mathcal{A}$ est isomorphe à l'algèbre des matrices carrées d'ordre p. Pour $\alpha \neq 0$, posons $F_\alpha = 1/(q - q^{-1}) e_\alpha$, et soit $\mathcal{H}_q$ le $\mathbf{Z}[q]$-sous-module de base $(F_\alpha)_{\alpha \neq 0}$. On a

$$[F_\alpha, F_\beta] = \frac{q^{<\alpha|\beta>} - q^{-<\alpha|\beta>}}{q - q^{-1}} F_{\alpha+\beta}.$$

Or $\frac{q^{<\alpha|\beta>} - q^{-<\alpha|\beta>}}{q-q^{-1}}$ appartient à $\mathbf{Z}[q]$, donc $\mathcal{H}_q$ est une $\mathbf{Z}[q]$-algèbre de Lie. Comme $\mathbf{C} \otimes \mathcal{H}_q$ est isomorphe a $\mathfrak{sl}(p)$, $\mathcal{H}_q$ est une $\mathbf{Z}[q]$-forme de $\mathfrak{sl}(p)$. Soit K un corps de caractéristique p. L'algèbre de Lie $\mathcal{H} = K \otimes \mathcal{H}_q$ est une algèbre de Lie de base $(E_\alpha)_{\alpha \neq 0}$ et le crochet est défini par $[E_\alpha, E_\beta] = < \alpha | \beta > E_{\alpha+\beta}$. Il est facile de voir que l'algèbre de Lie $\mathcal{H}$ est simple, mais n'est pas restreinte.

Le fait d'avoir trouvé une $\mathbf{Z}[q]$-forme à la place d'une $\mathbf{Z}$-forme est sans importance, mais cela a rendu les calculs plus faciles. En améliorant la construction précédente, on peut trouver une $\mathbf{Z}$-forme de $\mathfrak{sl}(p)$ dont la réduction modulo p est $\mathcal{H}$. Par exemple

choisissons une base $(v_n)_{0\leq n<p}$ de $\mathbf{C}^p$ et posons $Y = \{(i,j) \neq (0,0) | 0 \leq n,m < p\}$. Pour tout $(n,m) \in Y$, notons $e_{i,j}$ l'endomorphisme de $\mathbf{C}^p$ défini par $e_{i,j}.v_n = (1+p)^{ni}/p\, v_{n+j}$ si $n+j<p$ et $e_{i,j}.v_n = (1+p)^{ni}/p\, v_{n+j-p}$ si $n+j \geq p$. Bien que la définition des endomorphismes $e_{i,j}$ fasse intervenir un dénominateur, il est facile de montrer que $\mathcal{H}_{\mathbf{Z}} = \oplus_{i,j\in Y} \mathbf{Z} e_{i,j}$ est stable par crochet. De plus $\mathbf{Q}\otimes\mathcal{H}_{\mathbf{Z}}$ est isomorphe a $\mathfrak{sl}(p,\mathbf{Q})$ et $K\otimes\mathcal{H}_{\mathbf{Z}}$ est isomorphe à $\mathcal{H}$.

En quelque sorte $\mathcal{H}_{\mathbf{Z}}$ est une meilleure forme en p que la forme naturelle $\mathfrak{sl}(p,\mathbf{Z})$, puisque $K\otimes\mathcal{H}_{\mathbf{Z}}$ est simple tandis que $K\otimes\mathfrak{sl}(p,\mathbf{Z})$ ne l'est pas. Pour l'idée de cette réduction, voir [**GM**].

2.3. La Z-forme canonique de $\mathfrak{g}_{\mathbf{C}}(I)$

Rappelons que $\mathfrak{g}_{\mathbf{C}}(I)$ est l'algèbre de Lie d'un groupe algébrique $G_{\mathbf{C}}$ et que Chevalley et Steinberg ont défini une **Z**-forme naturelle $G_{\mathbf{Z}}$ de ce groupe (voir [**Stb**], [**T1-3**]). Appelons **Z**-*forme canonique* de $\mathfrak{g}_{\mathbf{C}}(I)$ l'anneau de Lie du groupe $G_{\mathbf{Z}}$, et notons-la $\mathfrak{g}_{\mathbf{Z}}^{can}(I)$. Nous allons maintenant comparer la forme canonique à la forme $\mathfrak{g}_{\mathbf{Z}}(I)$ définie à l'alinéa précédent.

Définir le groupe $G_{\mathbf{Z}}$ signifie contruire son anneau de fonctions $\mathbf{Z}[G_{\mathbf{Z}}]$ et le munir d'une structure d'anneau de Hopf. La définition de $\mathbf{Z}[G_{\mathbf{Z}}]$ est la suivante : considérons le sous-anneau $U_{\mathbf{Z}} \subset U(\mathfrak{g}_{\mathbf{C}}(I))$ engendré par les éléments $e_i^{(n)}$ et $f_i^{(n)}$ pour tout $i \in I$ et $n \geq 1$ (la notation $x^{(n)}$ signifie $\frac{1}{n!}x^n$). Comme $U(\mathfrak{g}_{\mathbf{C}}(I))$ peut être identifié aux distributions supportées à l'origine, il y a un couplage naturel $U(\mathfrak{g}(I)) \times \mathbf{C}[G_{\mathbf{C}}] \to \mathbf{C}$, $(u,f) \mapsto < u|f >$. On définit $\mathbf{Z}[G_{\mathbf{Z}}]$ comme l'ensemble des fonctions $f \in \mathbf{C}[G_{\mathbf{C}}]$ telles que $< u|f > \in \mathbf{Z}$ pour tout $u \in U_{\mathbf{Z}}$, et on vérifie que $\mathbf{Z}[G_{\mathbf{Z}}]$ est un sous-anneau de Hopf et qu'il s'agit d'une **Z**-forme de $\mathbf{C}[G_{\mathbf{C}}]$. Il est alors facile de montrer le fait suivant :

LEMME 4.— *L'anneau de Lie $\mathfrak{g}_{\mathbf{Z}}^{can}(I)$ est le sous-anneau de Lie de $\mathfrak{g}_{\mathbf{C}}(I)$ engendré par les éléments $\frac{1}{n!}ad(f_j)^n(f_i)$ et $\frac{1}{n!}ad(e_j)^n(e_i)$ pour tous $i,j \in I$ et $n \geq 0$.*

En utilisant les relations de Serre, on voit que :
(i) Si I est de type ADE, $\mathfrak{g}_{\mathbf{Z}}(I) = \mathfrak{g}_{\mathbf{Z}}^{can}(I)$,
(ii) Si I est de type BDF, notons a,b l'unique paire d'éléments de I tels que $a_{a,b} = -2$. Alors $\mathfrak{g}_{\mathbf{Z}}^{can}(I)$ est engendrée par $\mathfrak{g}_{\mathbf{Z}}(I)$ et les deux éléments $\frac{1}{2}ad(e_a)^2(e_b)$ et $\frac{1}{2}ad(f_a)^2(f_b)$. En particulier, $\mathfrak{g}_{\mathbf{Z}}(I)$ et $\mathfrak{g}_{\mathbf{Z}}^{can}(I)$ coïncident à la 2-torsion près.
(iii) Si I est de type G_2, posons $I = \{a,b\}$ de sorte que $a_{a,b} = -2$. Alors $\mathfrak{g}_{\mathbf{Z}}^{can}(I)$ est engendrée par $\mathfrak{g}_{\mathbf{Z}}(I)$ et les quatre éléments $\frac{1}{2}ad(e_a)^2(e_b)$, $\frac{1}{6}ad(e_a)^3(e_b)$, $\frac{1}{2}ad(f_a)^2(f_b)$, $\frac{1}{6}ad(f_a)^3(f_b)$. En particulier, $\mathfrak{g}_{\mathbf{Z}}(I)$ et $\mathfrak{g}_{\mathbf{Z}}^{can}(I)$ coïncident à la 6-torsion près.

Pour tout corps K, posons $\mathfrak{g}_K^{can}(I) = K \otimes \mathfrak{g}_{\mathbf{Z}}^{can}(I)$. On notera encore par h_i, e_i, f_i les images de ces éléments dans $\mathfrak{g}_K^{can}(I)$, et on pose $\mathfrak{h} = \oplus_{i \in I} K.h_i$ Notons $\mathfrak{z}$ l'ensemble des $h \in \mathfrak{h}$ tels que $[h, e_i] = 0$ pour tout i. Un élément $h = \sum_i x_i h_i$ appartient à $\mathfrak{z}$ si et seulement si $(x_i)_{i \in I}$ est dans le noyau de A mod. p. On a $det(A) = 1$ si I est de type G_2, F_4 ou E_8, $det(A) = 2$ si I est de type BC ou E_7, $det(A) = 3$ si I est de type E_6, $det(A) = 4$ si I est de type D et $det(A) = l + 1$ si I est de type A_l. On en déduit que $\mathfrak{z} = 0$, sauf dans l'un des cas suivants :

cas où $\mathfrak{z} = K$:

si I est de type A_{np-1}, ou

si $p = 3$ et I est de type E_6, ou

si $p = 2$ et I est de type B_l, C_l, D_{2n+1}, E_7.

cas où $\mathfrak{z} = K^2$:

si $p = 2$ et I est de type D_{2n}.

LEMME 5.— *Supposons que I ne soit pas de type A_1, B_l, C_l ou F_4 si $p = 2$ et que I ne soit pas de type G_2 si $p = 3$. Alors l'algèbre de Lie $\mathfrak{g}_K^{can}(I)/\mathfrak{z}$ est simple et restreinte.*

Preuve : Posons $\mathfrak{g} = \mathfrak{g}_K^{can}(I)/\mathfrak{z}$. Notons $(\alpha_i)_{i \in I}$ une base de $\mathbf{Z}^l$. L'algèbre de Lie admet une Z^l-graduation naturelle, pour laquelle e_i est de degré α_i, et f_i est de degré $-\alpha_i$ et cette Z^l-graduation induit une décomposition :

$$\mathfrak{g} = \mathfrak{h}/\mathfrak{z} \oplus_{\alpha \in \Delta} \mathfrak{g}_\alpha,$$

où $\mathfrak{g}_\alpha$ est la composante de degré $\alpha \in Z^l$ et $\Delta \subset Z^l$ est l'ensemble $\alpha \in Z^l \setminus 0$ tels que $\mathfrak{g}_\alpha \neq 0$. Supposons que $\mathfrak{g}$ ne soit pas simple et possède un idéal non nul J. La Z^l-graduation de $\mathfrak{g}$ induit l'action d'un tore $H = (K^*)^l$ dans le groupe des automorphismes de $\mathfrak{g}$. Dans la fermeture de l'orbite de J dans la grassmanienne de $\mathfrak{g}$, on trouvera un idéal de même dimension et stable par H. Donc on peut supposer que J est un idéal Z^l-gradué.

Posons $J_0 = J \cap \mathfrak{h}/\mathfrak{z}$. On veut montrer que J_0 contient au moins l'un des h_i, et on commence par montrer que J_0 est non nul. On utilise alors que, pour tout $\alpha \in \Delta$, l'algèbre de Lie engendrée par $\mathfrak{g}_{\pm\alpha}$ est isomorphe à $\mathfrak{sl}(2)$. Ce fait utilise que c'est le cas pour $\alpha = \alpha_i$ et l'action du groupe de Weyl que nous ne définirons pas (par définition, $\mathfrak{g}_K^{can}(I)$ est un G_K-module). Pour montrer que $J_0 \neq 0$, on peut supposer que J n'est pas réduit à J_0, donc J contient l'un des $\mathfrak{g}_\alpha$ et donc J_0 contient $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$, ce qui montre $J_0 \neq 0$. Soit $h \in J_0 \setminus 0$. Par définition de $\mathfrak{z}$, il existe $i \in I$ tel que $[h, e_i] \neq 0$. Donc e_i et f_i appartiennent à I et h_i appartient à J_0.

Maintenant on montre que J_0 contient h_i, pour tout $i \in I$. Pour cela on montre que si i et j sont connectés dans le diagramme de Dynkin, et si $h_i \in I_0$, alors h_j y appartient aussi par le même argument qu'avant, car $[h_i, e_j] \neq 0$ (les hypothèses assurant que $a_{i,j} \neq 0 \mod. p$). Comme le diagramme de Dynkin I est connexe, J_0 contient tous les h_i.

Il n'est pas difficile de montrer que $\mathfrak{g} = \mathfrak{h}/\mathfrak{z} \oplus [\mathfrak{h}, \mathfrak{g}]$, d'où $J = \mathfrak{g}$, et $\mathfrak{g}$ est donc simple.

3. ALGÈBRES DE TYPE CARTAN RESTREINTES

Dans cette section, nous allons d'abord définir les algèbres de type Cartan restreintes. Grâce au formalisme du calcul différentiel, nous décrirons en détail leur série dérivée et expliquerons pourquoi elles sont restreintes. Pour une approche basée sur des calculs plus explicites, on pourra se reporter aux références en bibliographie. En ce qui concerne la terminologie, il serait plus agréable d'appeler "algèbres de Cartan" ces algèbres, mais cela pourrait entraîner une confusion avec la notion classique de sous-algèbres de Cartan. "Algèbres de type Cartan" est la traduction de la terminologie anglaise "Cartan type algebras".

3.1. Définition des algèbres de type Cartan restreintes

Soit V un espace vectoriel de dimension n. Notons $\overline{S}V$ le quotient de SV par l'idéal engendré par v^p pour $v \in V$. Si $x_1, \ldots x_n$ est une base de V, alors

$$\overline{S}V = K[x_1, \ldots, x_n]/(\sum_{1 \leq i \leq n} K[x_1, \ldots, x_n]x_i^p).$$

L'algèbre $\overline{S}V$ admet pour base les monômes $x_1^{m_1} \ldots x_1^{m_n}$ où $m_i < p$ pour tout i. En particulier sa dimension est p^n. Il sera commode d'utiliser la terminologie de la géométrie algébrique, et de considérer $X = Spec\,\overline{S}V$ comme une variété et de noter par $K[X] = \overline{S}V$ l'espace des fonctions, par Ω_X^* l'espace des formes sur X et par W_X l'algèbre de Lie des champs de vecteurs sur X (i.e. $W_X = Der\,K[X]$). Bien que l'algèbre $K[X]$ contienne beaucoup d'éléments nilpotents, le calcul différentiel sur X est très similaire au calcul différentiel ordinaire sur les variétés lisses. En particulier, Ω_X^* est un $K[X]$-module libre : plus précisément $\Omega_X^* \simeq K[X] \otimes \wedge V$. Pour $\xi \in W_X$ et $\beta \in \Omega_X^*$, on notera par $i_\xi \beta$ le produit intérieur par ξ et par $\xi.\beta$ l'action naturelle de ξ sur β, c'est-à-dire la dérivée de Lie. La formule usuelle $\xi.\beta = i_\xi d\beta + d_\beta \xi$ est également valable dans ce cadre.

En fait Ω_X^* a une structure plus riche qu'une structure d'algèbre : l'espace des formes paires Ω_X^{2*} a une structure d'algèbre à puissances divisées (voir une définition formelle en section 5). Pour toute forme ω de degré $2k > 0$, sa puissance divisée $\omega^{(m)} = \omega^n/m!$ est bien définie. On a $d\omega^{(m)} = (d\omega) \wedge \omega^{(m-1)}$, $\xi.\omega^{(m)} = (\xi.\omega) \wedge \omega^{(m-1)}$ et $i_\xi \omega^{(m)} = (i_\xi \omega) \wedge \omega^{(m-1)}$, pour tout $\xi \in W_X$. Comme nous n'utiliserons que des puissances divisées de 2-formes, nous n'allons en donner une définition que pour celles-ci. Si ω est la 2-forme $1/2 \sum_{i,j} a_{i,j} dx_i \wedge dx_j$ avec $a_{i,j} + a_{j,i} = 0$, alors $\omega^{(m)} = \sum_{k_1<k_2\ldots<k_m} \mathrm{Pf}(a_{k_i,k_j}) dx_{k_1} \wedge \ldots \wedge dx_{k_n}$, où Pf désigne le pfaffien.

Introduisons les algèbres de Lie $W(n, \mathbf{1})$, $S(n, \mathbf{1})$, $H(2m, \mathbf{1})$ et $K(2m + 1, \mathbf{1})$. Par définition, $W(n, \mathbf{1})$ est l'algèbre de Lie W_X. Il sera commode d'utiliser concurremment ces deux notations, et de penser à $W(n, \mathbf{1})$ comme une classe d'isomorphisme d'algèbre de Lie. Pour toute forme β, on note $W_X(\beta)$ l'algèbre de Lie : $W_X(\beta) = \{\xi \in W_X | \xi.\beta = 0\}$, c'est-à-dire l'agèbre de Lie des champs de vecteurs qui préservent β. La *forme volume standard* est la n-forme $v_0 = dx_1 \wedge \ldots \wedge dx_n$ sur X. Par définition $S(n, \mathbf{1})$ est l'algèbre de Lie $W_X(v_0)$. Lorsque $n = 2m$, il sera parfois plus pratique pour décrire une forme symplectique de noter x_1, y_1, $x_2, \ldots, y_m$ la base de V. La *forme symplectique standard* est la 2-forme $\omega_0 = \sum_{1 \leq i \leq n} dx_i \wedge dy_i$. De même l'algèbre de Lie $W_X(\omega_0)$ sera notée $H(2m, \mathbf{1})$. Lorsque $n = 2m+1$, on notera parfois $z, x_1, y_1, x_2, \ldots y_n$ une base de V et on appelle la 1-forme $\alpha_0 = dz + 1/2 \sum_{1 \leq i \leq n} x_i \wedge dy_i - y_i dx_i$ la *forme de contact standard.* On pose alors $K(n, \mathbf{1}) = \{\xi \in W(n, \mathbf{1}) | \alpha_0 \wedge \xi.\alpha_0 = 0\}$. En fait $K(n, \mathbf{1})$ est l'algèbre de Lie des champs de vecteurs qui préservent la structure de contact définie par α_0, car la condition $\alpha_0 \wedge \xi.\alpha_0 = 0$ signifie que $\xi.\alpha_0 = f\alpha_0$ pour un certain $f \in K[X]$.

La signification du symbole $\mathbf{1}$, que nous appellerons le *niveau* dans les notations $W(n, \mathbf{1})$,..., $K(n, \mathbf{1})$ sera expliquée dans la section 6. À la différence des algèbres de Lie classiques, il n'y a pas d'isomorphismes exceptionnels entre les algèbres des quatre séries W, S, H et K (sauf l'évident $S(2, \mathbf{1}) = H(2, \mathbf{1})$). Rappelons que pour une algèbre de Lie L, $L^{(k)}$ désigne la série dérivée et il sera commode de poser $L^{(\infty)} = \cap_{k \geq 1} L^{(k)}$. Bien sûr, puisque les algèbres de Lie sont tacitement présumées de dimension finie, on a $L^{(\infty)} = L^{(k)}$ pour $k \ll 0$.

PROPOSITION 6.— *Les algèbres $W(n, \mathbf{1})^{(\infty)}$ (pour $n \geq 1$), $S(n, \mathbf{1})^{(\infty)}$ (pour $n \geq 2$), $H(2m, \mathbf{1})^{(\infty)}$ (pour $m \geq 1$) et $K(2m + 1, \mathbf{1})^{(\infty)}$ (pour $m \geq 1$) sont simples et restreintes.*

DÉFINITION 7.— *Les algèbres $W(n, \mathbf{1})^{(\infty)}$, $S(n, \mathbf{1})^{(\infty)}$ $H(n, \mathbf{1})^{(\infty)}$ et $K(n, \mathbf{1})^{(\infty)}$ (avec les mêmes conditions pour n et m) sont appelées les algèbres de type Cartan*

restreintes.

Nous allons maintenant décrire de manière plus détaillée les algèbres de type Cartan. Notons d'abord que ces algèbres de Lie possèdent une graduation naturelle. Pour fixer les conventions, rappelons qu'une graduation d'une algèbre de Lie L est une décomposition de L : $L = \oplus_{r\in\mathbf{Z}} L_r$ avec $[L_r, L_s] \subset L_{r+s}$. Comme on a $\overline{S}V = \oplus_{m\geq 0}\overline{S}^r V$, où $\overline{S}^r V$ est le sous-espace engendré par les monômes de degré r, l'algèbre commutative $\overline{S}V$ est graduée et cette graduation induit une graduation de l'algèbre de Lie $W(n, \mathbf{1})$, pour laquelle le champ de vecteurs $x_1^{m_1} \ldots x_1^{m_1}\partial/\partial x_i$ est homogène de degré $-1+m_1 \ldots +m_n$. On a donc $W(n, \mathbf{1}) = \oplus_{r\geq -1} W(n, \mathbf{1})_r$, où $W(n, \mathbf{1})_m = \overline{S}^{m+1}V \otimes V^*$. Comme la forme volume standard v_0 et la forme symplectique standard ω_0 sont homogènes, les algèbres de Lie $H(2m, \mathbf{1})$ et $H(2n, \mathbf{1})$ sont naturellement graduées. En revanche la forme de contact standard α_0 n'est pas homogène relativement à cette graduation. Lorsque $n = 2m + 1$, on considère la graduation de l'algèbre $\overline{S}V$ telle que les générateurs x_i, y_i soient homogènes de degré 1 et z homogène de degré deux. Relativement à cette graduation, α_0 est homogène et cela induit une graduation de $K(2m + 1, \mathbf{1})$ et l'on a $K(2m + 1, \mathbf{1}) = \oplus_{r\geq -2} K(2m + 1, \mathbf{1})_r$.

3.2. Un critère de simplicité

Soit $G = \oplus_{-1\leq s\leq r} G_s$ une algèbre de Lie graduée. On notera que G_{-1} est abélienne, G_0 est une sous-algèbre de Lie et chaque composante G_s est un G_0-module.

LEMME 8.— *L'algèbre de Lie G est simple dès que $r > 0$ et que les quatre hypothèses suivantes sont satisfaites :*

(i) *le commutant de G_{-1} dans G est réduit à G_{-1} ;*

(ii) *le G_0-module G_{-1} est simple ;*

(iii) *comme G_{-1}-module, G est engendrée par G_r ;*

(iv) *on a $[G_0, G_r] = G_r$.*

Preuve : Soit I un idéal non nul de G. Puisque $ad(G_{-1})$ agit de manière nilpotente, I contient un élément non nul x avec $ad(G_{-1})(x) = 0$. Par l'hypothèse (i), I coupe non trivialement G_{-1}. Par l'hypothèse (ii), I contient G_{-1}. L'hypothèse (iii) signifie que $[G_{-1}, G_{s+1}] = G_s$ pour tout $s < r$. Donc I contient toutes les composantes G_s avec $s < r$. Comme $r > 0$, I contient G_0, et par l'hypothèse (iv), I contient G_r. Donc $I = G$ et G est simple.

3.3. L'algèbre de Lie $W(n, \mathbf{1})$

PROPOSITION 9.— *L'algèbre de Lie $W(n, \mathbf{1})$ est de dimension np^n et elle est simple.*

L'algèbre de Lie $W(n, \mathbf{1})$ est souvent appelée *algèbre de Jacobson et Witt.*

Preuve : Identifions V^* aux champs de vecteurs à coefficients constants, i.e. de la forme $\sum_{1 \leq i \leq n} a_i \partial/\partial x_i$, où $a_i \in K$. Puisque $W(n, \mathbf{1}) = \overline{S}V \otimes V^*$, sa dimension est bien np^n. Pour prouver la simplicité de $W(n, \mathbf{1})$, nous allons vérifier qu'elle satisfait aux conditions du lemme 8. Notons que $W(n, \mathbf{1})_{-1} = V^*$.

La composante de plus haut degré $\overline{S}^{top}V$ de $\overline{S}V$ est $\overline{S}^{(p-1)n}V$. Il est facile de voir que le $U(V^*)$-module $\overline{S}V$ est engendré par $\overline{S}^{top}V$, et que seules les fonctions constantes sont annulées par V^*. Puisque $W(n, \mathbf{1}) = \overline{S}V \otimes V^*$, les conditions (i) et (iii) du lemme 8 sont satisfaites. Notons aussi que le degré maximal r de $W(n, \mathbf{1})$ est $n(p-1) - 1 \geq 1$.

Notons que $W(n, \mathbf{1})_0 = V \otimes V^* = \mathfrak{gl}(V)$. Comme $\mathfrak{gl}(V)$-modules, on a $W(n, \mathbf{1})_{-1} = V^*$ et $W(n, \mathbf{1})_1 \simeq (\wedge^n V)^{\otimes p-1} \otimes V^*$. Puisque ces deux $\mathfrak{gl}(V)$-modules sont simples et non triviaux, les conditions (ii) et (iv) sont aussi vérifiées. C.Q.F.D.

Remarque : Puisque la preuve n'utilise que l'inégalité $r = n(p-1) - 1 \geq 1$, $W(n, \mathbf{1})$ est aussi simple en caractéristique 3, et même en caractéristique 2 si $n \geq 2$.

3.4. La cohomologie de De Rham de X

Soit $\mathbf{m}$ l'idéal maximal de X. Bien que l'algèbre $K[X]$ soit locale, sa cohomologie de De Rham $H^*_{DR}(X)$ n'est pas triviale. Il existe un morphisme naturel d'anneau, l'opérateur de Cartier, $Car : \Omega^*_X \to H^*_{DR}(X)$ tel que $Car(f) = [f^p]$ et $Car(df) = [f^{p-1}df]$ pour tout $f \in K[X]$ (pour toute forme fermée β, nous notons par $[\beta]$ sa classe de cohomologie). À la différence du cas des variétés lisses, ici Car n'est pas un isomorphisme. Il est facile de montrer que :

LEMME 10.— *L'opérateur de Cartier induit un isomorphisme d'anneaux* :

$$C : \Omega^*_X/(\mathbf{m}\Omega^*_X) \to H^*_{DR}(X).$$

De manière explicite, $H^*_{DR}(X)$ est l'algèbre extérieure sur les générateurs x_i^{p-1} dx_i. En particulier, $\dim H^r_{DR}(X) = \binom{n}{r}$. Plus intrinsèquement, notons V^F l'espace vectoriel dont le groupe sous-jacent est V et où l'action de $\lambda \in K$ est la multiplication par λ^p. On a $H^*_{DR}(X) = \bigwedge V^F$.

Notons que $H^*_{DR}(X)$ est un W_X-module trivial, car la dérivée de Lie par ξ vaut $[d, i_\xi]$. En revanche, bien que le groupe $Aut(X)$ des automorphismes de X soit connexe, il agit non trivialement sur $H^*_{DR}(X)$. Par exemple pour tout $g = (g_{i,j}) \in GL(V)$, l'action de g sur V^F est déterminée par la matrice $(g^p_{i,j})$.

3.5. Un critère de "restreignabilité"

Pour toute forme $\beta \in \Omega_X$, l'algèbre de Lie $W_X(\beta)$ des champs de vecteurs qui préservent β est clairement restreinte. Lorsque β est fermée, la condition $\xi.\beta = 0$ est équivalente à la condition $i_\xi\beta$ est fermée. On dit alors que ξ *préserve strictement* β si $i_\xi\beta$ est exacte et l'on note $W'_X(\beta)$ l'ensemble des champs de vecteurs qui préservent strictement β. Par définition, l'application $W_X(\beta)/W'_X(\beta) \to H^*_{DR}(X)$, $\xi \mapsto [i_\xi\beta]$ est injective. Comme $H^*_{DR}(X)$ est un W_X-module trivial, on a $W'_X(\beta) \supset W_X(\beta)^{(1)}$.

Même si $W_X(\beta)$ est toujours restreinte, $W'_X(\beta)$ n'est pas toujours restreinte. Cependant, c'est le cas si β est exacte :

LEMME 11.— *Supposons que β soit exacte. Alors ξ^p appartient à $W'_X(\beta)$, pour tout $\xi \in W_X(\beta)$. En particulier, $W'_X(\beta)$ est une sous-algèbre de Lie restreinte de $W_X(\beta)$.*

Preuve : Choisissons une forme γ telle que $d\gamma = \beta$. On a $d\xi.\gamma = \xi.\beta = 0$. Donc $\xi^2.\gamma = di_\xi\xi.\gamma$ est exacte, et donc $\xi^p.\gamma$ est exacte. Or $\xi^p.\gamma = i_{\xi^p}\beta + di_{\xi^p}\gamma$, donc $i_{\xi^p}\beta$ est exacte, i.e. ξ appartient à $W'_X(\beta)$. C.Q.F.D.

3.6. L'algèbre de Lie $S(n,\mathbf{1})$, pour $n \geq 3$

Soit v_0 la forme volume standard de X. On note $\int_X : K[X] \to H^n_{DR}(X) \simeq K$ la forme linéaire $f \mapsto [f.v_0]$. Comme son noyau est l'analogue des fonctions de moyenne nulle, on le note $K_0[X]$. En fait f appartient à $K_0[X]$ si et seulement si elle ne contient pas le monôme $x_1^{p-1}\dots x_n^{p-1}$. La divergence (relative à v_0) d'un champ de vecteurs $\xi \in W_X$ est définie par $(\operatorname{div}\xi)v_0 = \xi.v_0$. Puisque $i_{W(n,\mathbf{1})}v_0 = \Omega_X^{n-1}$, une fonction f est la divergence d'un champ de vecteurs si et seulement si $\int_X f = 0$.

PROPOSITION 12 ($n \geq 3$).— (i) *L'algèbre de Lie $S(n,\mathbf{1})$ est de dimension* $1+(n-1)p^n$.

(ii) *On a $W'_X(v_0) = S(n,\mathbf{1})^{(1)}$. Cette algèbre de Lie est simple, restreinte et de dimension $(n-1)(p^n-1)$.*

Les algèbres de Lie de cette série sont appelées *spéciales*. L'algèbre de Lie $S(1,\mathbf{1})$ est de dimension 1. L'algèbre de Lie $S(2,\mathbf{1})$, isomorphe à $H(2,\mathbf{1})$, sera étudiée dans la suite avec les algèbres $H(2m,\mathbf{1})$.

Preuve : On calcule $\dim S(n,\mathbf{1})$ à l'aide de la suite exacte :

$$0 \to S(n,\mathbf{1}) \to W(n,\mathbf{1}) \to K_0[X] \to 0,$$

où la seconde application est la divergence. Puisque $i_{W(n,\mathbf{1})}v_0 = \Omega_X^{n-1}$, on a aussi une suite exacte :

$$0 \to W'_X(v_0) \to W_X(v_0) \to H^{n-1}_{DR}(X) \to 0,$$

où la seconde application est $\xi \mapsto [i_\xi v_0]$. On en déduit la dimension de $W'_X(v_0)$, et cette algèbre de Lie est restreinte par le lemme 11.

Il reste à montrer la simplicité de $W'_X(v_0)$. On pose $G = W'_X(v_0)$ et on veut vérifier que G satisfait aux quatre hypothèses du lemme 8. On a $G_{-1} = V^*$ et $G_0 = \mathfrak{sl}(V)$. L'assertion (i) a déjà été prouvée pour l'algèbre W_X, et l'assertion (ii) est triviale. Vérifions (iii) et (iv). L'application $\xi \mapsto i_\xi v_0$ identifie W'_X aux $(n-1)$-formes exactes. Il existe donc un morphisme naturel surjectif de W'_X-modules : $\psi : \Omega_X^{n-2} \to W'_X$. Or on a $\Omega_X^{n-2} \simeq \overline{S}V \otimes \bigwedge^2 V^*$ comme $G_0 \oplus G_{-1}$-module. Comme le G_1-module Ω_X^{n-2} est engendré par sa composante de plus haut degré, il en est de même de $W'_X(v_0)$, et le degré de cette composante est $r = n(p-1) - 2$. Ceci vérifie l'assertion (iii). Enfin le G_0-module G_r est isomorphe au G_0-module simple et non trivial $\bigwedge^2 V^*$, ce qui montre l'assertion (iv) (c'est ici que l'hypothèse $n \neq 2$ est utilisée).

Remarque : Puisque la preuve n'utilise que l'inégalité $r = n(p-1) - 2 \geq 1$, $S(n, \mathbf{1})$ est simple en toutes caractéristiques.

3.7. Algèbres de Poisson

On appelle *algèbre de Poisson* toute algèbre commutative unitaire A muni d'une structure d'algèbre de Lie (dont le crochet est noté $\{,\}$) qui satisfasse à l'identité de Leibniz $\{fg, h\} = f\{g, h\} + g\{f, h\}$.

On suppose maintenant que $n = 2m$, on note $x_1, y_1, x_2, \ldots$ une base de V et on rappelle que $\omega_0 = \sum_{1 \leq i \leq m} dx_i \wedge dy_i$ est la forme symplectique standard. Il existe une structure naturelle de structure de Poisson sur $K[X]$, et le crochet $\{f, g\}$ de deux fonctions f, g, appelé *crochet de Poisson*, est défini par l'identité $df \wedge dg \wedge \omega_0^{(m-1)} = \{f, g\}\omega_0^{(m)}$ (ce qui est bien défini car $\omega_0^{(m)}$ est la forme volume standard sur X). De manière explicite, on a :

$$\{f, g\} = \sum_{1 \leq i \leq m} f_{x_i} g_{y_i} - g_{x_i} f_{y_i},$$

où la notation h_x signifie $\partial/\partial x h$.

PROPOSITION 13.— *On a $K[X]^{(1)} = K_0[X]$ et l'algèbre de Lie $K_0[X]/K$ est simple (et de dimension $p^{2m} - 2$).*

Preuve : Pour toute fonction f, notons H_f le champ de vecteurs tel que $H_f.g = \{f, g\}$ pour tout $g \in K[X]$. On peut aussi définir H_f par l'identité $df = i_{H_f}\omega$, donc H_f préserve la forme symplectique ω_0. Comme H_f préserve la forme volume $v_0 = \omega_0^{(m)}$,

on a $H_f.K[X] \subset K_0[X]$ et on a donc $K[X]^{(1)} \subset K_0[X]$. Comme $\{x_1, y_1\} = 1$, on a aussi $K \subset K[X]^{(1)}$. Donc l'egalité $K[X]^{(1)} = K_0[X]$ résultera de la simplicité de $K_0[X]/K$, ce que nous allons maintenant prouver.

Posons $G = K_0[X]/K$. Il existe une graduation naturelle de G telle que $G_s = K_0[X]_{s+2}$ pour tout $s \geq -1$. Lorsque nous identifions G à une sous-algèbre de W_X, G_0 s'identifie à $\mathfrak{sp}(V)$ et $G_{-1} \simeq V$ aux champs à coefficients constants. Pour prouver la simplicité de G, nous allons vérifier qu'elle satisfait aux quatre assertions du lemme 8.

L'assertion (i) a déjà été vérifiée pour W_X, et l'assertion (ii) est triviale. On a déjà vérifié que le plus grand degré de $K[X]$ est $m(p-1)$ et que l'on a $G_{-1}.K[X]_s = K[X]_{s-1}$ pour tout $s \leq m(p-1)$. Comme $K_0[X] = \oplus_{s<m(p-1)} K[X]_s$, on obtient $G_{-1}.K_0[X]_s = K_0[X]_{s-1}$ pour tout $s < m(p-1)$. Par conséquent le plus grand degré r de G est $2m(p-1)-3$, et l'on a $[G_{-1}, G_{s+1}] = G_s$ pour tout $s < r$. Ceci vérifie l'assertion (iii). Comme G_r est isomorphe à V comme G_0-module, l'assertion (iv) du lemme 8 est vérifiée. Donc G est simple. C.Q.F.D.

Remarque : Puisque la preuve n'utilise que l'inégalité $2m(p-1)-3 \geq 1$, $K_0[X]/K$ est aussi simple en caractéristique 3, et même en caractéristique 2 si $2m \geq 4$.

Nous voulons aussi établir que $K_0[X]/K$ est restreinte, ou de manière équivalente que $K_0[X]$ est "restreignable". La stratégie que l'on va suivre est de se ramener au cas $m = 1$.

Le produit tensoriel $A \otimes B$ de deux algèbres de Poisson A et B est muni d'une structure d'algèbre de Poisson. Le crochet de Poisson de $A \otimes B$ est défini par $\{a \otimes b, a' \otimes b'\} = \{a, a'\} \otimes bb' + aa' \otimes \{b, b'\}$, pour tous $a, a' \in A$, $b, b' \in B$. Nous dirons qu'une algèbre de Poisson A est *bonne* si et seulement si :

(i) On a $x^p \in \{A, A\}$ pour tout $x \in A$.

(ii) Pour tout $x \in A$, il existe $y \in \{A, A\}$ tel que $ad(x)^p = ad(y)$,

LEMME 14.— *Soient A, B deux algèbres de Poisson.*

(i) *On a $\{A \otimes B, A \otimes B\} = \{A, A\} \otimes B + A \otimes \{B, B\}$.*

(ii) *Si A et B sont bonnes, alors $A \otimes B$ est bonne.*

Preuve : Il est clair que $\{A \otimes B, A \otimes B\} \subset \{A, A\} \otimes B + A \otimes \{B, B\}$. Par ailleurs l'algèbre de Lie dérivée de $A \otimes B$ contient $\{A \otimes B, A \otimes K\} = \{A, A\} \otimes B$ ainsi que $\{A \otimes B, K \otimes B\} = A \otimes \{B, B\}$, d'où l'égalité.

Supposons que A et B soient bonnes. Il est clair que, pour tout $x \in A \otimes B$, x^p appartient à $\{A, A\} \otimes \{B, B\}$ qui est inclus dans $\{A \otimes B, A \otimes B\}$. Il s'agit maintenant de vérifier que :

(*) pour tout $x \in A \otimes B$, il existe $y \in \{A \otimes B, A \otimes B\}$ tel que $ad(x)^p = ad(y)$. Par l'identité de Jacobson (théorème 1), on a $ad(x_1 + x_2)^p = ad(x_1)^p + ad(x_2)^p + ad(J_p(x_1+x_2))$. Or $J_p(x_1+x_2)$ est un commutateur, donc $J_p(x_1+x_2) \in \{A\otimes B, A\otimes B\}$ pour tous x_1, $x_2 \in A\otimes B$. Donc il suffit de vérifier (*) pour un système de générateurs de l'espace vectoriel $A\otimes B$. En particulier, il suffit de le vérifier pour les x de la forme $a \otimes b$.

Soient $a \in A$, $X \in \{A, A\}$, $b \in B$ et $Y \in \{B, B\}$ tels que $ad(a)^p = ad(X)$, $ad(b)^p = ad(Y)$. On a $ad(a \otimes b) = M_a \otimes ad(b) + ad(a) \otimes M_b$ où M_a, M_b désignent la multiplication par a, b. Comme les quatre opérateurs M_a, M_b, $ad(a)$ et $ad(b)$ commutent, on a :

$$ad(a \otimes b)^p = M_{a^p} \otimes ad(b)^p + ad(a)^p \otimes M_b^p.$$

Or $ad(a^p) = ad(b^p) = 0$. D'où :

$$ad(a \otimes b)^p = ad(X \otimes b^p + a^p \otimes Y).$$

Par hypothèse, $X \otimes b^p + a^p \otimes Y$ est bien dans $\{A \otimes B, A \otimes B\}$, ce qui prouve le lemme 14.

PROPOSITION 15.— *Pour tout $x \in K[X]$, il existe $y \in K_0[X]$ tel que $ad(x)^p = ad(y)$. En particulier $K_0[X]/K$ est restreinte.*

Preuve : Par le lemme 14, on se ramène au cas $m = 1$. Rappelons que $K[X] = \overline{S}V$. Pour tout $k < p$, $\overline{S}^k V$ est linéairement engendré par les éléments de la forme v^k où $v \in V$. Donc il existe une base B de $K[X]$ tel que tout élément $f \in B$ est de l'un des deux types suivants :

(i) $f = v^k$ pour un certain $v \in V$, ou

(ii) f est homogène de degré ≥ 4.

En utilisant l'identité de Jacobson, il suffit de prouver que $ad(f)^p$ est intérieure pour tout $f \in B$. En fait, on a $ad(f)^p = 0$ pour tout $f \in B$. Dans le premier cas, $f = v^k$, on a $H_f = kv^{k-1}H_v$. Comme $H_v.v = 0$, on a $H_f^p = k^p v^{p(k-1)} H_v^k = 0$ car $H_v^p = 0$. Dans le second, $ad(f)$ est un opérateur homogène de degré ≥ 2, donc $ad(f)^p$ est homogène de degré $\geq 2p$. Or la différence entre le deux degré de $K[X]$ est bornée par $2(p-1)$, donc $ad(f)^p = 0$.

Remarque : La preuve a utilisé $p > 3$, mais le résultat est aussi vrai pour $p = 2, 3$.

3.8. L'algèbre de Lie $H(2m, \mathbf{1})$

Soit ω_0 la forme symplectique standard sur X. Par définition de $W_X'(\omega_0)$, pour tout $\xi \in W_X'(\omega_0)$, il existe $f \in K[X]$ avec $i_{xi}\omega = df$. Notons $W_X''(\omega_0)$ l'ensemble des $\xi \in W_X'(\omega_0)$ tel que $i_{xi}\omega = df$ avec $\int_X f = 0$. Il est clair que $W_X''(\omega_0)$ est un idéal de $W_X(\omega_0)$. Rappelons la notation $H(2m, \mathbf{1}) = W_X(\omega)$.

PROPOSITION 16.— (i) *Les algèbres de Lie* $W_X''(\omega_0) \subset W_X'(\omega_0) \subset H(2m, \mathbf{1})$ *sont restreintes, et de dimension* $p^{2m} - 2$, $p^{2m} - 1$ *et* $p^{2m} + 2m - 1$.

(ii) *Si* $m = 1$, $H(2m, \mathbf{1})^{(1)} = W_X'(\omega_0)$, $H(2m, \mathbf{1})^{(2)} = W_X''(\omega_0)$ *et* $H(2m, \mathbf{1})^{(2)}$ *est simple.*

(iii) *Si* $m > 1$, $H(2m, \mathbf{1})^{(1)} = W_X''(\omega_0)$ *et* $H(2m, \mathbf{1})^{(1)}$ *est simple.*

Preuve : Reprenons les notations et résultats de l'alinéa précédent. L'application $f \in K[X] \mapsto H_f$ induit un isomorphisme $K_0[X]/K \simeq W_X''(\omega_0)$. Par les propositions 13 et 15, l'algèbre de Lie W_X'' est simple, restreinte et de dimension $p^{2m} - 2$. Par le lemme 8, W_X' est restreinte. Les dimensions de $W_X'(\omega_0)$ et $W_X(\omega_0)$ se calculent en utilisant que $W_X''(\omega_0)$ est un idéal de codimension 1 de $W_X'(\omega_0)$ et que l'application $W_X(\omega_0)/W_X'(\omega_0) \to H^1_{DR}(X), \xi \mapsto [i_\xi \omega_0]$ est un isomorphisme. Il ne reste plus qu'à calculer la série dérivée de $H(2m, \mathbf{1})$, ou ce qui revient au même la série dérivée de $W_X(\omega_0)/W_X''(\omega_0)$.

Considérons d'abord le cas $m \neq 1$. Notons que les isomorphismes

$$W_X'(\omega_0)/W_X''(\omega_0) \simeq K[X]/K_0[X] \text{ et } W_X(\omega_0)/W_X'(\omega_0) \simeq H^1_{DR}(X)$$

sont des isomorphismes d'espaces vectoriels gradués, à un décalage par 2 près (car ω_0 est homogène de degré 2). Donc $W_X(\omega_0)/W_X''(\omega_0)$ est une algèbre de Lie graduée, et ses deux composantes sont de degré $2m(p-1) - 2$ et $p - 2$. Pour qu'une algèbre de Lie graduée $L = L_a \oplus L_b$ avec $0 < a < b$ ne soit pas commutative, il est nécessaire que $2a = b$. Donc si $m \neq 1$, $W_X(\omega_0)/W_X''(\omega_0)$ est commutative, et l'on a donc $H(2m, \mathbf{1})^{(1)} = W_X''(\omega_0)$.

Pour $m = 1$, il est nécessaire de faire un calcul direct. Une base de $W_X(\omega_0)/W_X''(\omega_0)$ est $x_1^{p-1}\partial/\partial y_1$, $y_1^{p-1}\partial/\partial x_1$ et $H_{x_1^{p-1}y_1^{p-1}}$, et il est facile de voir que

$$[x_1^{p-1}\partial/\partial y_1, y_1^{p-1}\partial/\partial x_1] = -H_{x_1^{p-1}y_1^{p-1}},$$

ce qui achève la preuve. C.Q.F.D.

Remarque : On établit facilement que, pour tout $\xi \in H(2m, \mathbf{1})$, on a $\xi^p \in H(2m, \mathbf{1})^{(\infty)}$ (c'est encore vrai pour $p = 3$, et même $p = 2$ si $m > 1$).

3.9. Un autre critère de simplicité

Pour prouver la simplicité des algèbres de contact, nous aurons besoin d'un critère légèrement différent. Soit $G = \oplus_{-2 \leq s \leq r} G_s$ une algèbre de Lie graduée. On pose $G^- = G_2 \oplus G_1$.

LEMME 17 ($p \geq 3$).— *Faisons les hypothèses suivantes :*

(i) *Le commutant de G^- dans G est réduit à G_{-2},*

(ii) $\dim G_{-2} = 1$,

(iii) *Comme G^--module, G est engendré par G_r,*

(iv) *Il existe $h \in [G_{-2}, G_2]$ tel que $ad(h)(x) = nx$ pour tout $x \in G_n$.*

Sous les hypothèses précédentes, on a alors :

(i) *Si $r \neq 0 \bmod p$, l'algèbre de Lie est simple.*

(ii) *Si $[G,G]_r = 0$, alors $[G,G] = \oplus_{-2 \leq s < r} G_s$ et $[G,G]$ est simple.*

Preuve : Supposons d'abord que $r \neq 0 \bmod p$, et soit I un idéal non nul de G. Comme $ad(G^-)$ agit de manière nilpotente, I contient un vecteur non nul x avec $ad(G^-)(x) = 0$. Par les hypothèses (i) et (ii), I contient G_{-2}. L'hypothèse (iv) implique que I contient h, et donc aussi $G_{-1} = [h, G_{-1}]$ et $G_r = [h, G_r]$. Comme I contient G^- et G_r, il contient le G^--module engendré par G_r. Par l'hypothèse (iii), $I = G$. Donc G est simple.

Supposons maintenant que $[G,G]_r = 0$, et posons $G' = \oplus_{s<r} G_s$. On a $G' \supset [G,G]$, donc G' est une sous-algèbre de Lie. Puisque $G_2 \neq 0$, on a $r \geq 2$. Comme $[h, G_r] = 0$, on a $r = 0 \bmod p$. En particulier $r \geq 3$ et $G_2 \subset G'$. Soit I un idéal non nul de G'. Par la même preuve qu'avant, on montre que $G^- \subset I$, $h \in I$ et puisque $[h, G_{r-1}] = G_{r-1}$ et $[h, G_{r-2}] = G_{r-2}$, on a aussi $G_{r-1} \oplus G_{r-2} \subset I$. Donc I contient le G^--module engendré par $G_{r-1} \oplus G_{r-2}$.

Or l'hypothèse (iii) implique que l'on a $G_s = [G_{-1}, G_{s+1}] + [G_{-2}, G_{s+2}]$ pour tout $s < r$. Donc G' est engendré par $G_{r-1} \oplus G_{r-2}$ comme G^--module, et on a donc $I = G'$. Donc G' est simple et $G' = [G,G]$. C.Q.F.D.

3.10. L'algèbre de Lie $K(2m+1, \mathbf{1})$

On suppose que $n = 2m+1$. Soit α_0 la forme de contact standard sur X. On rappelle que $K(2m+1, \mathbf{1}) = \{\xi \in W(2m+1, \mathbf{1}) | \alpha_0 \wedge \xi.\alpha_0 = 0\}$. On notera $K_0(2m+1, \mathbf{1})$ l'ensemble des champs de contacts $\xi \in K(2m+1, \mathbf{1})$ tels que $\int_X i_\xi \alpha_0 = 0$. En général, $K_0(2m+1, \mathbf{1})$ n'est pas une sous-algèbre de Lie.

PROPOSITION 18.— (i) *L'algèbre de Lie restreinte $K(2m+1, \mathbf{1})$ est de dimension p^{2m+1}, et lorsque $m+2 \neq 0$ (mod. p), elle est simple.*

(ii) *Si $m+2$ est divisible par p, alors on a $K(2m+1,\mathbf{1})^{(1)} = K_0(2m+1,\mathbf{1})$. De plus $K(2m+1,\mathbf{1})^{(1)}$ est simple, restreinte et de dimension $p^{2m+1}-1$.*

Avant de prouver la proposition 18, nous allons décrire plus explicitement la structure d'algèbre de Lie de $K(2m+1,\mathbf{1})$.

LEMME 19.— (i) *L'application $I : K(2m,\mathbf{1}) \to K[X], \xi \mapsto i_\xi\alpha_0$ est bijective.*

(ii) *Lorsque $m+2 = 0 \bmod p$, $K_0(2m+1,\mathbf{1})$ est un idéal de codimension 1.*

Preuve : Comme α_0 est une forme de contact, W_X se décompose en somme de deux $K[X]$-modules : $W_X = \alpha_0^\perp \oplus K[X].E$, où $\alpha_0^\perp$ l'ensemble des champs de vecteurs ξ tels que $i_\xi\alpha_0 = 0$ et où E est le *champ d'Euler*, i.e. l'unique champ de vecteurs tel que $i_E\alpha_0 = 1$ et $i_E d\alpha_0 = 0$. Tout champ de vecteurs ξ se décompose de manière unique en $\xi = \xi_1 + \xi_2$, où $x_1 \in \alpha_0^\perp$ et ξ_2 est proportionnel à E. De même toute 1-forme β s'écrit de manière unique comme $\beta = i_{\xi_1}d\alpha_0 + g\alpha_0$, où $\xi_1 \in \alpha_0^\perp$ et $g \in K[X]$.

Prouvons l'injectivité de I. Soit $\xi = \xi_1 + \xi_2 \in K(2m+1,\mathbf{1})$ avec $I(\xi) = 0$. Donc par hypothèse $\xi_2 = 0$. De plus $\xi.\alpha_0 = i_{\xi_1}d\alpha_0$. Or $\xi.\alpha_0$ est proportionnel à α_0. Donc $\xi_2 = 0$, d'où $\xi = 0$.

Prouvons maintenant la surjectivité de I. Soit $f \in K[X]$. Posons $\xi_2 = f.E$. Il existe un unique $\xi_1 \in \alpha_0^\perp$ et une unique $g \in K[X]$ tels que : $df = i_{\xi_1}d\alpha_0 + g\alpha_0$. Posons $\xi = -\xi_1 + \xi_2$. On a $\xi.\alpha = i_{x_1}d\alpha_0 + di_{x_2}\alpha_0 = i_{x_1}d\alpha_0 + df = -g\alpha_0$, donc ξ appartient à $K(2m+1,\mathbf{1})$. De plus $I(\xi) = f$, d'où la surjectivité de I.

Vérifions maintenant l'assertion (ii). Soit $\xi \in K(2m,\mathbf{1})$. Puisque $\alpha_0 \wedge \xi.\alpha_0 = 0$, il existe une fonction, notée $\mathrm{Div}(\xi)$ telle que $\xi.\alpha_0 = \mathrm{Div}(\xi)\alpha_0$. Comparons sa divergence de contact $\mathrm{Div}(\xi)$ à sa divergence ordinaire. On notera que $\xi.d\alpha_0 = d\xi.\alpha_0 = Div(\xi)d\alpha_0 - \alpha_0 dDiv(\xi)$. Or $\alpha_0 \wedge (d\alpha_0)^{(m)}$ est la forme volume standard, d'où :

$$\begin{aligned}\xi.v_0 &= (\xi.\alpha_0) \wedge (d\alpha_0)^{(m)} + \alpha_0 \wedge (\xi d\alpha_0) \wedge (d\alpha_0)^{(m-1)} \\ &= \mathrm{Div}(\xi).v + \alpha_0 \wedge d\mathrm{Div}(\xi)\alpha_0 \wedge (d\alpha_0)^{(m-1)} \\ &= (m+1)\mathrm{Div}(\xi).\end{aligned}$$

Donc $(m+1)\mathrm{Div}(\xi) = \mathrm{div}(\xi)$.

Pour $\xi, \eta \in K(2m+1,\mathbf{1})$, on a :

$$\begin{aligned}\xi.I(\eta) &= \xi.i_\eta\alpha_0 \\ &= i_{[\xi,\eta]}\alpha_0 + i_\eta\xi.\alpha_0 \\ &= I([\xi,\eta]) + \mathrm{Div}(\xi)I(\eta).\end{aligned}$$

D'où $I([\xi,\eta]) = \xi.I(\eta) - \mathrm{Div}(\xi)I(\eta)$.

Donc lorsque $m+2$ est divisible par p, on a : $\mathrm{Div}\,\xi = -\mathrm{div}\,\xi$, et on obtient :

$$I([\xi,\eta]) = \mathrm{div}\,\theta(\eta)\xi,$$

ou de manière équivalente $K(2m+1,\mathbf{1})^{(1)} \subset K_0(2m+1,\mathbf{1})$, ce qui prouve l'assertion (ii) (en notant que tout champ de vecteurs est combinaison linéaire de champs de la forme $f\zeta$, où $f \in K[X]$ et $\zeta \in K(2m+1,\mathbf{1})$, on obtient même $K(2m+1,\mathbf{1})^{(1)} = K_0(2m+1,\mathbf{1})$).

Remarque : Le calcul précédent montre que lorsque $m+1$ est divisible par p, tout champ de contact préserve la forme volume.

La bijection $I : K(2m+1,\mathbf{1}) \to K[X]$ induit une structure d'algèbre de Lie sur $K[X]$ et on notera par $\{f,g\}$ le crochet correspondant. Un calcul direct montre que :

$$\{f,g\} = \sum_{1\le i\le m} f_{x_i}g_{y_i} - g_{x_i}f_{y_i} + g_z\Big(-2f + \sum_{1\le i\le m} x_if_{x_i} + y_if_{y_i}\Big) - f_z\Big(-2g + \sum_{1\le i\le m} x_ig_{x_i} + y_ig_{y_i}\Big).$$

À la différence du crochet de Poisson, le crochet de contact n'est pas une bidérivation, puisqu'il contient aussi des termes linéaires. Notons aussi que I n'est pas un morphisme de $K(2m+1,\mathbf{1})$-modules, puisque l'on a $\xi.I(\eta) = I(\xi), I(\eta) + I(\eta)\mathrm{Div}\,\xi$ pour tout $\xi,\eta \in K(2m+1,\mathbf{1})$.

Preuve de la proposition 18 : Nous allons appliquer la proposition 17 pour vérifier la simplicité de $K(2m+1,\mathbf{1})$ (si $m+2 \not\equiv 0 \bmod. p$) ou de $K_0(2m+1,\mathbf{1})$ (dans le cas contraire). Posons $G = K[X]$. Rappelons que G est une algèbre de Lie graduée, et l'isomorphisme $I : K(2m+1,\mathbf{1}) \to K[X]$ respecte la graduation à un décalage de 2 près. On $G_{-2} = K$, $G_{-1} = \oplus_{1\le i\le m} K.x_i \oplus K.y_i$ et la composante de degré maximale est $G_r = K.z^{p-1}x_1^{p-1}y_1^{p-1}\ldots$, où $r = (2m+1)(p-1)-2$.

L'élément 1 de G_{-2} agit comme $-2\partial/\partial z$, et les éléments x_i et y_i comme $\partial/\partial y_i - x_i\partial/\partial z$ et $-\partial/\partial x_i - y_i\partial/\partial z$. Donc il est clair que G est engendré par G_r comme G^--module, et que le commutant de G^- est G_{-2}. Les hypothèses (i), (ii), (iii) du lemme 16 sont donc vérifiées.

Vérifions maintenant l'hypothèse (iv). Posons $h = -z$. Comme h agit comme $-2 + 2z\partial/\partial z + \sum_{1\le i\le m} x_i\partial/\partial x_i + y_i\partial/\partial y_i$, h agit sur G_n comme n. De plus $4h = \{1, z^2\}$, donc $h \in [G_{-2}, G_2]$, donc h satisfait à l'hypothèse (iv).

Si $m+2 \not\equiv 0 \bmod. p$, alors $r \not\equiv 0 \bmod. p$ et par le lemme 17 l'algèbre de Lie $K(2m+1,\mathbf{1})$ est simple.

Supposons maintenant $m+2 = 0 \bmod. p$. Par le lemme 19, on a $[G,G] \subset K_0(2m+1,\mathbf{1})$ et par définition on a $K_0(2m+1,\mathbf{1}) = \oplus_{s<r} G_s$. À nouveau le lemme 17 nous permet de conclure que $K_0(2m+1,\mathbf{1})$ est simple.

4. LE THÉORÈME DE BLOCK ET DE WILSON

Rappelons que K désigne un corps algébriquement clos de caractéristique 0. En sections 2 et 3, nous avons décrit certaines algèbres de Lie simples restreintes. Le théorème qui suit a été conjecturé par Kostrikin et Shafarevitch [**KS**] en 1966 et prouvé par Block et Wilson en 1988 [**BW1**][**BW3**].

THÉORÈME 20 (Block et Wilson).— *Supposons $p > 7$. Alors toute algèbre de Lie simple restreinte est isomorphe à une algèbre classique ou à une algèbre de type Cartan restreinte.*

Le cas $p = 11$ présente déjà des difficultés spéciales (mais résolues). Les experts conjecturent que le résultat de Block et Wilson est valable aussi en caractéristique 7, mais les difficultés techniques ne sont pas toutes résolues pour $p = 7$. Pour $p = 5$, ce résultat ne peut se généraliser, car il y a une algèbre de Lie restreinte simple de dimension 125 qui n'est ni classique ni de type Cartan : c'est l'algèbre de Lie de Melikian, voir [**Me**], [**St7**]. Les experts conjecturent qu'en caractéristique 5, il n'y a pas d'autres exceptions. À ce jour, il n' y a pas de conjecture raisonnable pour $p = 3$ et $p = 2$.

5. DÉFAUT AU LEMME DE DARBOUX ET PREMIERS EXEMPLES D'ALGÈBRES DE LIE NON RESTREINTES

Pour illustrer la complexité de la classification des algèbres de Lie simples non restreintes, nous allons d'abord décrire un problème voisin, c'est-à-dire la classification des formes volumes et symplectiques sur X. Une version plus générale et plus technique de cette section fera l'objet d'une publication séparée. On peut aussi comparer les conclusions de cette section avec [**BGOSW**].

5.1 Notations pour la section 5

On reprendra les notations de la section 3. En particulier $K[X]$ désigne l'anneau $K[x_1,\ldots,x_n]/(x_1^p,\ldots,x_n^p)$. Pour $r \geq 0$, notons $K[X]_{(r)}$ le sous-espace engendré par les monômes de degré $\geq r$, i.e. $K[X]_{(r)}$ est l'ensemble des fonctions qui s'annulent

au moins r fois à l'origine. On notera $\mathbf{m} = K[X]_{(1)}$ l'idéal maximal de $K[X]$ et $V = \mathbf{m}/\mathbf{m}^2$. Pour toute forme fermée β, on note $W_X(\beta)$ l'algèbre de Lie des champs de vecteurs ξ tels que $\xi.\beta = 0$, et $W'_X(\beta)$ la sous-algèbre des champs de vecteurs tels que $i_\xi\beta$ soit exacte.

5.2. Lemme de Darboux

Notons $Aut(X)$ le groupe des automorphismes de la variété X. Un automorphisme de X est de la forme $(x_1, x_2, \ldots) \mapsto A(x) + \phi(x)$, où $A \in GL(n)$ et où $\phi(x) = (\phi_1(x), \phi_2(x), \ldots)$, avec $\phi_1, \phi_2, \ldots \in K[X]_2$. Le groupe $Aut(X)$ est un groupe algébrique connexe d'algèbre de Lie $\oplus_{r\geq 0} W(n, \mathbf{1})_r$. Son radical unipotent $Aut_1(X)$ est l'ensemble des automorphismes de la forme $(x_1, x_2, \ldots) \mapsto x + \phi(x)$ et l'on a $Aut(X)$, produit semi-direct de $GL(V)$ par $Aut_1(X)$.

Le groupe $Aut(X)$ agit naturellement sur l'espace des formes Ω^*_X, respecte l'idéal maximal m et commute à d. Il agit donc naturellement sur Ω^*_X/m et sur la cohomologie de X. Bien que $Aut(X)$ soit connexe, son action sur $H^*_{DR}(X)$ est non triviale (c'est un phénomème typique de la caractéristique finie : son action infinitésimale sur $H^*_{DR}(X)$ est triviale). En revanche, le groupe $Aut_1(X)$ agit trivialement sur $H^*_{DR}(X)$. Par le lemme 10, il existe un isomorphisme naturel d'anneaux, $C : \Omega^*_X/m \to H^*_{DR}(X)$ (il ne s'agit pas d'un morphisme d'algèbres, car il n'est pas linéaire). Par conséquent le groupe $GL(V) = Aut(X)/Aut_1(X)$ agit sur deux copies de $\bigwedge V$, via les isomorphismes $\bigwedge V = \Omega^*_X/m$ et sur $\bigwedge V = C^{-1}H^*_{DR}(X)$. Ces deux actions sont les actions naturelles de $GL(V)$ sur $\bigwedge V$.

Bien que l'algèbre $K[X]$ soit locale, elle ne satisfait pas au lemme de Darboux : il existe des formes volumes et des formes symplectiques (lorsque $n = 2m$) qui ne sont pas conjuguées aux formes standard. En fait les formes standards sont exactes, et il existe des formes volumes ou symplectiques qui ne soient pas exactes.

Soit β une forme fermée. Notons $\psi(\beta) \in \bigwedge V \oplus \bigwedge V$ le couple (β_1, β_2) tel que $\beta_1 = \beta$ modulo $\mathbf{m}\Omega^*_X$ et $\beta_2 = C^{-1}[\beta]$. Il est clair que si deux formes fermées β, β' sont conjuguées par $Aut(X)$, alors $\psi(\beta)$ et $\psi(\beta')$ sont conjuguées sous $GL(V)$.

LEMME 21.— *Soient β, β' deux formes volumes ou deux formes symplectiques sur X. Alors β et β' sont conjuguées par $Aut(X)$ si et seulement si $\psi(\beta)$ et $\psi(\beta')$ sont conjuguées sous $GL(V)$.*

Preuve : Posons $r = 2$ si β est une forme symplectique, et $r = n$ si β est une forme volume. Supposons que $\psi(\beta)$ et $\psi(\beta')$ soient conjugués par $GL(V)$, et montrons que β et β' sont conjuguées. Identifions Ω^r_X à $K[X] \otimes \bigwedge^r V$. Décomposons β et β' en

composantes homogènes : $\beta = \sum_{t\geq 0} \beta_t$ et $\beta' = \sum_{t\geq 0} \beta'_t$, où $\beta_t, \beta'_t \in K[X]_t \otimes \bigwedge^r V$. Quitte à conjuguer par $GL(V)$, on peut supposer que $\beta_0 = \beta'_0$ et que $\beta' - \beta$ est exacte.

Soit $s \geq 1$ un entier tel que $\beta'_t = \beta_t$ pour tout $t < s$. On veut montrer qu'il existe un conjugué β'' de β tel que $\beta''_t = \beta_t$ pour tout $t < s+1$ et $\beta'' - \beta$ soit encore exacte. Par définition d'une forme volume ou symplectique, $i_{W_X}\beta_0 = \Omega_X^{r-1}$, donc $W_X.\beta_0 = d\Omega_X^{r-1}$. Il existe donc un champ de vecteurs $\xi \sum_{1\leq i\leq n} f_i\partial/\partial x_i$ tel que $\xi.\beta_0 = \beta_s - \beta'_s$. On peut de plus supposer que chaque f_i est homogène de degré $s+1$. Soit Φ l'automorphisme $(x_1, x_2, \ldots) \mapsto (x_1 + f_1, x_2 + f_2, \ldots)$ et posons $\beta'' = \phi\beta'$. On vérifie facilement que β'' coïncide avec β jusqu'à l'ordre s. Comme $\Phi \in Aut_1(X)$, il agit trivialement sur $H^*_{DR}(X)$ et $\beta'' - \beta$ est encore exacte.

En répétant cette procédure, et en notant que toute composante de degré $> (p-1)n$ est nulle, on conjugue β' et β par un automorphisme de X.

5.3. Remarques sur les algèbres filtrées

Une *algèbre de Lie filtrée* est une algèbre de Lie L avec une filtration décroissante $L = L_{(1)} \supset L_{(0)} \subset L_{(1)} \ldots$ telle que $[L_r, L_s] \subset L_{r+s}$. Le gradué associé à L, i.e. l'espace vectoriel ${}^{gr}L = \oplus_{r\geq -1} L_{(r)}/L_{(r+1)}$, est muni d'une structure d'algèbre de Lie. Il est facile de voir que L est une déformation de ${}^{gr}L$. Par conséquent toute propriété satisfaisant à un principe de semi-continuité et vérifiée pour ${}^{gr}L$ est aussi vérifiée pour L : par exemple si ${}^{gr}L$ est simple, ou si son centre est trivial, ou si ${}^{grL} = [{}^{gr}L, {}^{gr}L]$, L satisfait à la même propriété. On a aussi : $({}^{gr}L)^{(n)} \subset {}^{gr}(L^{(n)})$, pour tout $n \geq 1$. Cependant, ${}^{gr}L$ peut être "restreignable" sans que L le soit (en revanche, si ${}^{gr}L$ a un centre trivial et si L est "restreignable", alors ${}^{gr}L$ est "restreignable"). Donnons un critère de simplicité d'une algèbre filtrée :

LEMME 22.— *Soit L une algèbre de Lie filtrée, et soit L' un idéal avec $L' \supset L^{(\infty)}$. Supposons que $({}^{gr}L)^{(\infty)}$ est simple, et de codimension N, et que l'une des hypothèses suivantes soit satisfaite :*

(i) *Si L' est de codimension N, ou*

(ii) *Si L' est de codimension $N-1$, de centre trivial et admet une forme bilinéaire invariante et non dégénérée.*

Alors $L' = L^{(\infty)}$ et $L^{(\infty)}$ est simple.

Preuve : On a ${}^{gr}L' \supset ({}^{gr}L)^{(\infty)}$. Sous l'hypothèse (i), on a donc $L' = L^{(\infty)}$ et $L^{(\infty)}$ est simple. Assumons maintenant l'hypothèse (ii). Soit $I \neq q0$, ${}^{gr}L'$ un idéal de ${}^{gr}L'$. Puisque $({}^{gr}L)^{(\infty)}$ est simple et de codimension 1 dans ${}^{gr}L'$ on a $I = ({}^{gr}L)^{(\infty)}$ (auquel cas I est de codimension 1), ou bien $I \cap ({}^{gr}L)^{(\infty)} = 0$ (auquel cas I est de dimension 1). Or L' ne peut contenir un idéal J de codimension 1 (car sinon $J \supset [L, L]$ et son

orthogonal est central), ni de dimension 1 (car son orthogonal serait de codimension 1). Donc L' est simple, et $L' = L^{(\infty)}$. C.Q.F.D.

L'algèbre de Lie $W(n, \mathbf{1})$ a une filtration naturelle :

$$W(n, \mathbf{1}) = W(n, \mathbf{1})_{(-1)} \supset W(n, \mathbf{1})_{(0)} \supset \ldots,$$

où $W(n, \mathbf{1})_{(r)}$ est l'ensemble des champs de vecteurs qui s'annulent au moins $r + 1$ fois à l'origine, i.e. à coefficients dans $K[X]_{(r+1)}$. Cette filtration induit aussi une filtration de $S(n, \mathbf{1})$ et de $H(2m, \mathbf{1})$ (lorsque $n = 2m$). Les algèbres de contact ont aussi une filtration, mais qui est définie par un procédé un peu différent.

5.4. Formes volumes sur X

Posons $v_1 = du_1/u_1 \wedge \ldots du_n/u_n$, où $u_i = 1 + x_i$. Il est facile de voir que $[v_1] \in H^n_{DR}(X)$ est non nulle. Par conséquent, v_1 n'est pas conjuguée sous $Aut(X)$ à la forme volume sandard v_0. D'après le lemme précédent toute forme volume sur X est conjuguée à v_0 ou à λv_1, où $\lambda \in K$. Par conséquent, il n'existe que deux algèbres de Lie spéciales sur X, à savoir $W_X(v_0) \simeq S(n, \mathbf{1})^1$ et $W_X(v_1)$.

L'algèbre $W_X(v_1)$ est restreinte. Comme v_1 n'est pas exacte, nous ne pouvons pas appliquer le lemme 8. En fait l'algèbre de Lie $W'_X(v_1)$ n'est pas restreinte comme le montre le lemme suivant.

LEMME 23 ($n \geq 2$).— *On a $W'_X(v_1) = W_X(v_1)^{(1)}$, et l'algèbre de Lie $W_X(v_1)^{(1)}$ est simple, non restreinte et de dimension $(n-1)p^{n-1}$.*

Preuve : Prouvons d'abord que $W'(v_1)$ n'est pas restreinte. Posons $\xi = u_2 u_1 \partial/\partial u_1$. On a $i_{xi} v' = du_2 \wedge du_3/u_3 \wedge \ldots du_n/u_n = d(u_2 \wedge du_3/u_3 \wedge \ldots du_n/u_n)$, donc $\xi \in W'(v_1)$. Notons que $u_1 \partial/\partial u_1 u_1^k = k u_1^k$. Donc $(u_1 \partial/\partial u_1)^p u_1^k = k^p u_1^k = k u_1^k$, donc $(u_1 \partial/\partial u_1)^p = u_1 \partial/\partial u_1$. D'où $\xi^p = u_2^p u_1 \partial/\partial u_1 = u_1 \partial/\partial u_1$. Or $i_{\xi^p} v' = du_2/u_2 \wedge du_3/u_3 \wedge \ldots du_n/u_n$ n'est pas exacte, et donc ξ^p n'appartient pas à $W'(v_1)$. Ainsi $W'(v_1)$ n'est pas restreinte (ni "restreignable", car son centre est trivial).

Prouvons maintenant la simplicité de $W'_X(v_1)$ pour $n \geq 3$. Posons $L = W_X(v_1)$, $L' = W'_X(v_1)$. Notons que $W_X(v_1)$ s'identifie à l'espace des $n-1$ formes fermées et $W'_X(v_1)$ s'identifie à l'espace des formes exactes. En particulier, $\dim L = \dim S(n, \mathbf{1})$ et $\dim L' = \dim S(n, \mathbf{1})^1$. On en déduit que ${}^{gr}W_X(v_1) = S(n, \mathbf{1})$. Par le lemme 22, on obtient que $L' = [L, L]$ et que L' est simple.

La simplicité de $W'_X(v_1)$ pour $n = 2$ sera prouvée dans la section suivante.

5.5. Formes symplectiques

Supposons que $n = 2m$, et notons $x_1, y_1, \ldots, y_m$ les générateurs de $K[X]$. En utilisant le lemme on peut complètement classifier les formes symplectiques sur X. Pour $k \leq m$, considérons l'ensemble $C_{m,k}$ des k-uplets $(\mu_1, \ldots, \mu_k)$ où $\mu_i = (\lambda_i, m_i)$ appartient à $K \times \mathbf{Z}_{>0}$ et $\sum_i m_i = m$. On considère comme égaux deux k-uplets qui ne diffèrent que par une permutation simultanée des indices. On posera $C_m = \cup_k C_{m,k}$. En fait C_m est l'ensemble qui paramétrise naturellement les classes de conjugaison de matrices de taille m. Pour $\mu = (\lambda, r) \in K \times \mathbf{Z}_{>0}$, on pose $\mathrm{rk}\,\mu = r$ si $\lambda \neq 0$ et $\mathrm{rk}\,\mu = r - 1$ sinon. Le *rang* $\mathrm{rk}\,\mu_*$ d'un uplet $\mu_* = (\mu_1, \ldots, \mu_k) \in C_m$ est l'entier $\sum_i \mathrm{rk}\,\mu_i$.

LEMME 24.— *Il existe une bijection naturelle $\mu : \omega \mapsto \mu(\omega)$ des classes de conjugaison de formes symplectiques sur C_n.*

De plus le rang de $\mu(\omega)$ est le plus grand entier r tel que $[\omega^{(r)}] \neq 0$.

Preuve : D'après le lemme 21, les formes symplectiques β sur X sont classifiées par les classes de conjugaison de paires $(\beta_0, \beta_1) \in \bigwedge^2 V \oplus \bigwedge^2 V$, où β_0 est la réduction modulo m de β et $\beta_1 = C^{-1}[\beta]$. Notons que β_0 est une forme symplectique de $\bigwedge^2$, et on peut la ramener à une forme normale, et l'on notera par $< \mid >$ la forme symplectique linéaire correspondante sur V. Par conséquent les classes de conjugaison de formes symplectiques de X sont en bijection avec l'ensemble des orbites $\bigwedge^2 V$ sous le groupe $SP(V)$.

Pour un espace linéaire symplectique W, appelons endomorphisme antisymétrique toute application linéaire $A : W \to W$ telle que la forme $w_1, w_2 \in W \mapsto < Aw_1 | w_2 >$ soit antisymétrique. Considérons l'ensemble $\mathcal{P}$ des paires (W, A) formées d'un espace linéaire symplectique $W \neq 0$ et d'un endomorphisme antisymétrique A. Pour (W_1, A_1), (W_2, A_2) dans $\mathcal{P}$, on définit leur somme directe $(W_1 \oplus W_2, A_1 \oplus A_2)$ de manière évidente. Une paire indécomposable est une paire qui n'est pas une somme directe. Les paires indécomposables (W, A) de dimension $2d$ sont faciles à classifier : l'endomorphisme A n'admet qu'une seule valeur propre λ, et la matrice nilpotente $N = A - \lambda$ a deux blocs de Jordan de taille d. On pose alors $\mu(W, A) = (\lambda, d)$.

Tout endomorphisme antisymétrique A de V se décompose en somme directe d'indécomposables $(V, A) = \oplus_{i \leq k}(W_i, A_i)$ et l'entier k et le k-uplet $(\mu(W_1, A_1), (\mu(W_2, A_2), \ldots)$ décrivent exactement la classe de conjugaison de A sous $SP(V)$.

Or $\bigwedge^2 V$ s'identifie aux endomorphismes antiymétriques de V, d'où le premier résultat. De plus $\beta^{(m)}$ n'est pas exacte si et seulement si β_1 est symplectique, i.e. aucun des λ_i n'est nul. C.Q.F.D.

Soit $\lambda \in K^*$. Pour $(\lambda, r) \in K \times \mathbf{Z}_{>0}$, on pose $\lambda.(\lambda, r) = (\lambda^{1/p}\lambda, r)$ et pour tout k-uplet $\mu = (\mu_1, \ldots, \mu_k)$ on pose $\lambda.\mu = ((\lambda.\mu_1, \ldots, \lambda.\mu_k)$. Notons $\omega \mapsto \mu(\omega)$ la bijection entre les classes de conjugaison de formes symplectiques qui sont en bijection avec l'ensemble C_n. On a $\mu(\lambda\omega) = \lambda\mu\omega$.

Or $W_X(\omega) = W_X(\lambda\omega)$, donc les algèbres $W_X(\omega)$ sont paramétrisées C_n/K^*, c'est-à-dire par une réunion d'espaces projectifs. Plus précisément, notons C_n^{gen} l'ensemble des m-uplets $(\mu_1, \ldots, \mu_m$ avec $\mu_i = (\lambda_i, 1)$ et $\lambda_i \neq 0$. Alors C_n^{gen}/K^* est isomorphe à KP^{m-1}, et $C_n \setminus C_n^{gen}/K^*$ est une réunion d'espaces projectifs de dimension $\leq m-2$. J'ignore si cette paramétrisation est bijective, mais il est très probable que les isomorphismes entre les algèbres de Lie correspondant à deux valeurs du paramètre soient exceptionnels. En bref on s'attend à une classification faisant apparaître $m-1$ paramètres.

Soit ω une forme symplectique sur X. Pour tout $f \in K[X]$, notons $\int_X f.\omega^{(m)}$ la classe de cohomologie de $f.\omega^{(m)}$, et notons $K_0^\omega[X]$ l'ensemble des fonctions $f \in K[X]$ telles que $\int_X f.\omega^{(m)}] = 0$, et $W_X''(\omega)$ l'ensemble des champs de vecteurs ξ tels que $i_\xi\omega = df$ pour un certain $f \in K_0^\omega[X]$. Il est clair que l'on a $W_X''(\omega) \subset W_X'(\omega) \subset W_X(\omega)$. On a toujours $W_X(\omega)/W_X'(\omega) = H^1_{DR}(X)$, donc $W_X'(\omega)$ est de codimension $2m$. De plus on a $W_X''(\omega) = W_X'(\omega)$ si $[\omega^{(m)}] \neq 0$. Sinon, $W_X''(\omega)$ est de codimension 1 dans $W_X'(\omega)$.

LEMME 25.— *Soit ω une forme symplectique sur X.*

(i) *On a $W_X(\omega)^{(\infty)} = W_X''(\omega)$, et cette algèbre est simple, mais n'est restreinte que si ω est conjugué à ω_0.*

(ii) *La dimension de $W_X(\omega)^{(\infty)}$ est $p^{2m}-2$ si $\omega^{(m)}$ est exacte, ou $p^{2m}-1$ sinon.*

(iii) *Si $\operatorname{rk}\mu(\omega) \neq m-1$, on a $W_X(\omega)^{(1)} = W_X''(\omega)$.*

(iv) Si $\operatorname{rk}\mu(\omega) \neq m-1$, on a $W_X(\omega)^{(1)} = W_X'(\omega)$ et $W_X(\omega)^{(2)} = W_X''(\omega)$.

Preuve : Nous n'allons que prouver les deux premières assertions. On a toujours $W_X(\omega)^{(1)} \subset W_X'(\omega)$ et $W_X(\omega)^{(2)} \subset W_X''(\omega)$. Donc l'identité $W_X(\omega)^{(\infty)} = W_X''(\omega)$ résultera de la simplicité de $W_X''(\omega)$ que nous allons maintenant prouver.

Notons que $K[X]$ a une structure d'algèbre de Poisson, et $K_0^\omega[X]$ est idéal relativement à sa structure d'algèbre de Lie.

Supposons d'abord que $\omega^{(m)}$ soit exacte. On a alors $K \subset K_0^\omega[X]$. Comme ${}^{gr}(K_0^\omega[X]/K) \simeq K_0[X]/K = H(2m, \mathbf{1}^{(\infty)}$, l'algèbre de Lie $K[X]/K$ est simple par la proposition 16 et le lemme 22. Comme $W_X''(\omega) \simeq K_0^\omega[X]/K$, cette algèbre de Lie est simple.

Supposons maintenant que $\omega^{(m)}$ ne soit pas exacte. Il est clair que l'application bilinéaire $f, g \in K[G] \mapsto \int_X f.\omega^{(m)} \in H^{2m}_{DR}(X) \simeq K$ est une forme bilinéaire non

dégénérée. Puisque $\int_X 1.\omega^{(m)} \neq 0$, la restriction de cette forme bilinéaire à $K_0^{\omega}[X]$ est non dégénérée. Posons $L = W_X(\omega)$ et $L' = W_X'(\omega) = W_X''(\omega)$. Puisque L' est isomorphe à $K_0^{\omega}[X]$, l'algèbre de Lie L' admet une forme bilinéaire invariante non dégénérée. Par la proposition 16, les algèbres filtrées L et L' satisfont aux hypothèses du lemme 22. Donc $L' = W_X''(\omega)$ est simple. C.Q.F.D.

5.6. Exemples d'algèbres hamiltoniennes non restreintes

Pour finir, nous allons décrire toutes les algèbres de Lie $W_X(\omega)^{(\infty)}$ telles que $\mu(\omega)$ soit de rang maximal, c'est-à-dire telles que $\omega^{(m)}$ ne soit pas exacte.

Posons $E = \mathbf{F_p}^{2m}$, et soit $B : E \times E \to K, \alpha, \beta \mapsto < \alpha|\beta >$ une forme bilinéaire alternée. On considère l'algèbre de Poisson $\mathcal{H}^B$ de base $(E_\alpha)_{\alpha \in E}$ et dont la structure de Poisson est donnée par $E_\alpha E_\beta = E_{\alpha+\beta}$, $E_\alpha, E_\beta = < \alpha|\beta > E_{\alpha+\beta}$.

Il est clair que le sous-espace $\mathcal{H}_0^B$ de base $(E_\alpha)_{\alpha \neq 0}$ est une sous-algèbre de Lie.

LEMME 26.— (i) *L'algèbre de Lie $\mathcal{H}_0^B$ est isomorphe à $W_X(\omega)^{(\infty)}$ pour une certaine forme symplectique ω telle que $[\omega^{(m)}] \neq 0$.*

(ii) *Pour toute forme symplectique ω sur X avec $[\omega^{(m)}] \neq 0$, il existe une forme bilinéaire alternée B telle que $W_X(\omega)^{(\infty)} \simeq \mathcal{H}_0^B$.*

L'assertion (i) est évidente. Nous n'allons pas prouver l'assertion (ii), mais simplement indiquer que cela résulte facilement du résultat suivant (théorème d'isogénie de Lang) : pour tout $M \in GL(m, K)$, il existe $N \in GL(m, K)$ avec $M.N = N^F$, où N^F désigne la matrice dont les coefficients sont les puissances p-ièmes des coefficients de N.

Le groupe $K^* \times GL(E)$ agit narurellement sur l'espace des formes alternées B : $E \times E \to K$ et, si B, B' sont conjuguées par ce groupe, les algèbres de Lie $\mathcal{H}_0^B$ et $\mathcal{H}_0^{B'}$ sont clairement isomorphes. Puisque le groupe $GL(E)$ est fini, les algèbres de Lie $\mathcal{H}_0^B$ sont paramétrées par une famille à $2m(2m-1)-1$ paramètres. Or les classes d'isomorphismes des algèbres $\mathcal{H}_0^B$ ne sont paramétrées que par $m-1$ paramètres. Donc il existe de nombreux isomorphismes non triviaux entre ces algèbres de Lie lorsque $2m \geq 4$.

6. ALGÈBRE DE TYPE CARTAN DE NIVEAU ARBITRAIRE

Nous avons déjà vu que l'on peut produire des algèbres simples non restreintes à partir de formes fermées non exactes. En fait les algèbres à puissance divisée fournissent aussi des algèbres de Lie non restreintes. La définition des algèbres de

type Cartan générales est apparue tout d'abord dans [**Kc3**], mais nous allons suivre celle de [**W1**].

6.1. Algèbres à puissances divisées

Soit A une K-algèbre commutative locale d'idéal maximal M. Une structure de puissance divisée est une suite d'applications $\gamma_r : M \to A, f \mapsto f^{(r)}$ indexées par les entiers $r > 0$ telles que :

(i) $f^{(0)} = 1$ et $f^{(r)} \in M$ pour $r > 0$.

(ii) $f^{(1)} = f$

(iii) $(f+g)^r = \sum_{a+b=r} f^a g^b$

(iv) $f^{(r)} f^{(s)} = \frac{(r+s)!}{r!s!} f^{(r+s)}$

(v) $(hf)^r = h^r f^{(r)}$

(vi) $(f^{(s)})^{(r)} = \frac{(rs)!}{r!(s!)^r} f^{rs}$,

pour tous $f, g \in M$, $h \in A$ $r \geq 0$ et $s > 0$.

Notons qu'une structure de puissance divisée est entièrement définie par la seule application $\gamma_p : f \mapsto p^{(p)}$. En effet, les axiomes impliquent que l'on a :

(vi) $f^{(r)} = \frac{1}{r!} f^r$ dès que $r < p$,

(vii) $f^{(p^j)} = (f^{(p^{j-1})})^{(p)}$,

(viii) $f^{(r)} = \prod_{j \geq 0} (f^{(p^j)})^{(r_j)}$, si $r = \sum_{j \geq 0} r_j p^j$ où $0 \leq r_j < p$,

et la dernière formule peut s'écrire comme $\gamma_{(r)} = \prod_{j \geq 0} \frac{1}{r_j!} \gamma_p^{(j)}(f^{r_j})$. Dans une algèbre à puissance divisée A, on a $f^p = 0$ pour tout $f \in M$.

Soit V un espace vectoriel de dimension finie. Puisque V est un groupe, l'algèbre des fonctions polynomiales sur V, à savoir SV^*, est aussi une cogèbre. Le coproduit $\Delta : SV^* \to SV^* \otimes SV^*$ est défini par $\Delta F(u,v) = F(u+v)$ pour tout $F \in SV^*$. Donc $(SV^*)^*$ a une structure d'algèbre. Sur un corps de caractéristique 0, cette algèbre est une algèbre de séries formelles, car on a alors $(S^n V^*)^* = S^n V$ pour tout entier n. En caractéristique finie p, ce n'est plus le cas car on ne peut identifier les espaces des invariants et des co-invariants pour l'action du groupe symétrique S_n sur $V^{\otimes n}$. C'est pourquoi nous noterons $(SV^*)^*$ par $K[\hat{X}]$ et nous dirons que $(SV^*)^*$ est l'ensemble des fonctions sur la variété $\hat{X}$. En fait la variété $\hat{X}$ ou, si l'on préfère l'algèbre $K[\hat{X}]$, est munie de structures supplémentaires. Tout d'abord, puisque $K[\hat{X}]$ est un dual, $K[\hat{X}]$ est une algèbre topologique. De plus $K[\hat{X}]$ est muni d'une structure de puissances divisées que nous allons maintenant définir.

Notons que l'opérateur différentiel $\frac{1}{n!}(d/dt)^n : K[t] \to K[t]$ est bien défini : en effet $\frac{1}{n!}(d/dt)^n t^m = \binom{m}{n} t^{m-n}$. Pour tout $v \in V$ et $n \geq 0$, on définit l'élément $u^{(n)}$ de

$K[\hat{X}]$ par

$$u^{(n)}F = \frac{1}{n!}(d/dt)^n F(tu)|_{t=0}.$$

Soit $M = \prod_{n\geq 1}(S^n V^*)^*$ l'idéal maximal de $K[\hat{X}]$. Il est clair que M est topologiquement engendré par les éléments $u^{(n)}$ pour $u \in V$ et $n \geq 1$. Si $n \geq 1$ n'est pas une puissance de p, on a $(x+y)^n \neq x^n + y_n \bmod. p$, i.e. il existe $a \neq 0, p$ tel que le coefficient de $x^a y^{n-a}$ dans le développement de $(x+y)^n$ soit non nul $\bmod. p$. Puisque $\binom{n}{a}u^{(n)} = u^{(a)}u^{(n-a)}$, on voit que M/M^2 est topologiquement engendré par les éléments $u^{(p^n)}$ pour $n \geq 0$ et $u \in V$.

L'application γ_p est donc uniquement définie par $\gamma_p(u^{(p^n)}) = u^{(p^{n+1})}$ si $u \in V$ et $\gamma_p(x) = 0$ si $x \in M^2$. Pour vérifier que γ_p est bien définie, il suffit d'observer que les applications $V \to M/M^2, u \mapsto u^{(p^n)}$ sont additives en u. L'application γ_p engendre une p-structure (pour laquelle la puissance divisée n-ième d'un élément u de V est bien $u^{(n)}$).

De manière explicite, choisissons une base $x_1, \ldots, x_n$ de V. Alors les éléments de $K[\hat{X}]$ s'écrivent comme des sommes formelles $\sum_\alpha a_\alpha x^{(\alpha)}$, où α parcourt l'ensemble des n-uplets $\alpha_1, \ldots, \ldots \alpha_n$ et $x^{(\alpha)} = x_1^{(\alpha_1)} \ldots x_n^{(\alpha_n)}$. Pour tout $r \geq 0$, on note $K[\hat{X}]_{(r)}$ les fonctions $f = \sum_\alpha a_\alpha x^{(\alpha)}$ telles que $a_\alpha = 0$ dès que $\alpha_1 + cdots + \alpha_n < r$. Un isomorphisme Φ de $\hat{X}$ est un isomorphisme continu d'algèbre $\Phi : K[\hat{X}] \to K[\hat{X}]$ tel que $\Phi(f^{(p)}) = \Phi(f)^{(p)}$ pour tout $f \in m$.

Par définition, un champ de vecteurs de $\hat{X}$ est une dérivation continue ξ de $K[\hat{X}]$ telle que $\xi.f^{(p)} = f^{(p-1)}\xi.f$ pour tout $f \in m$. L'ensemble $W_{\hat{X}}$ des champs de vecteurs de $\hat{X}$ est une algèbre de Lie. Définissons l'opérateur continu $\partial/\partial x_i$ par la formule $\partial/\partial x_i\, x^{(\alpha)} = x_1^{(\alpha_1 - 1)} \ldots x_n^{(\alpha_n)}$. Alors $\partial/\partial x_i$ appartient à $W_{\hat{X}}$ et tout élément ξ de $W_{\hat{X}}$ est de la forme $\sum_{1\leq i\leq n} f_i \partial/\partial x_i$. Pour tout $r \geq 0$, on note $(W_{\hat{X}})_{(r)}$ l'espace des champs de vecteurs $\xi = \sum_{1\leq i\leq n} f_i \partial/\partial x_i$ avec $f_i \in K[\hat{X}]_{(r)}$. Notons tout de suite que $W_{\hat{X}}$ n'est pas restreinte. En effet, on a $\xi^p.f^{(p)} = f^{(p-1)}\xi^p.f + (\xi.f)^p$. En revanche la sous-algèbre $(W_{\hat{X}})_{(1)}$ est restreinte, et c'est l'algèbre de Lie du groupe pro-algébrique $Aut(\hat{X})$.

L'algèbre $\Omega^*_{\hat{X}}$ des formes de $\hat{X}$ est la $K[\hat{X}]$-algèbre topologique engendrée par les symboles $\mathrm{d}f$, pour $f \in K[\hat{X}]$ et soumise aux relations usuelles :

$\mathrm{d}fg = f\mathrm{d}g + g\mathrm{d}f$

$\mathrm{d}f \wedge \mathrm{d}f = 0,$

pour tous $f, g \in K[\hat{X}]$, ainsi qu'à la relation additionnelle :

$\mathrm{d}f^{(p)} = f^{(p-1)}\mathrm{d}f$, pour tout $f \in M$.

Comme pour une variété ordinaire, les opérations usuelles du calcul différentiel sur $\Omega^*_{\hat{X}}$ (dérivée de Lie, contractions, différentielle extérieure d) sont bien définies, mais pas l'opérateur de Cartier.

Si l'on note encore par $H^*_{DR}(\hat{X})$ la cohomologie du complexe $(\Omega^*_{\hat{X}}, d)$, on voit que la cohomologie de $\hat{X}$ est triviale : $H^0_{DR}(\hat{X}) = K$ et $H^i_{DR}(\hat{X}) = 0$ pour $i > 0$. Par conséquent, il n'y a pas d'obstruction au lemme de Darboux pour $\hat{X}$.

6.2. Algèbre de type Cartan de dimension infinie

L'algèbre de Lie $W_{\hat{X}}$ est notée $W(n)$. L'algèbre de Lie $W(n)$ admet une filtration $W(n) = W(n)_{(-1)} \supset W(n)_{(0)} \supset W(n)_{(1)} \ldots$, où $W(n)_{(i)}$ est le sous-espace topologiquement engendré par les champs de vecteurs $x^\alpha \partial/\partial x_i$ où $\alpha_1 + \alpha_2 + \ldots \geq i + 1$. Soit $v_0 = dx_1 \wedge \ldots \wedge dx_n$ la forme volume standard sur $\hat{X}$. L'algèbre de Lie $W_{\hat{X}}(v) = \{\xi \in W_{\hat{X}} | \xi.\alpha = 0\}$ sera aussi notée $S(n)$.

Lorsque $n = 2m$, on notera $x_1, y_1, x_2, \ldots y_n$ une base de V et soit $\omega = \sum_{1 \leq i \leq n} dx_i \wedge dy_i$ la forme symplectique standard. L'algèbre de Lie $W_{\hat{X}}(\omega) = \{\xi \in W_{\hat{X}} | \xi.\omega = 0\}$ sera notée $H(2m)$.

Lorsque $n = 2m+1$, on note $z, x_1, y_1, x_2, \ldots y_n$ une base de V et on pose $K(2m+1) = \{\xi \in W_{\hat{X}}) | \alpha \wedge \xi.\alpha = 0\}$, où $\alpha = dz + 1/2 \sum_{1 \leq i \leq n} x_i \wedge dy_i - y_i dx_i$ est la forme de contact standard. De même, la condition $\alpha \wedge \xi.\alpha = 0$ signifie que $\xi.\alpha = f\alpha$ pour un certain $f \in K[\hat{X}]$.

Il est facile de montrer que les algèbres de dimension infinie $W(n)$, $S(n)$, $H(2m)$ et $K(2m+1)$ sont simples. Ce sont les *algèbres de type Cartan* de dimension infinie.

6.3. Définition des algèbres de type Cartan

Soit $\mathbf{m} = (m_1, \ldots, m_n)$ un n-uplet d'entiers positifs. On notera $F[n, \mathbf{m}]$ la sous-algèbre de dimension finie de $K[\hat{X}]$ engendrée par les monômes x^α avec $\alpha_i < p^{m_i}$. On pose aussi $W(n, \mathbf{m}) = \oplus_{1 \leq i \leq n} F[n, \mathbf{m}] \partial/\partial x_i$. Il est clair que $W(n, \mathbf{m})$ est une sous-algèbre de Lie de dimension $np^{m_1 + m_2 \cdots}$.

Soit Φ un isomorphisme de $\hat{X}$. Notons que Φ agit naturellement sur $W_{\hat{X}}$ et pour toute sous-algèbre de Lie L de $W_{\hat{X}}$ on note L^Φ sa conjuguée par Φ. On définit quatre séries d'algèbres de Lie $Y(n, \mathbf{m}, \Phi)$ (ici Y est un symbole qui désigne W, S, H ou K) par :

$$Y(n, \mathbf{m}, \Phi) = Y(n) \cap W(n, \mathbf{m})^\Phi.$$

Pour un automorphisme Φ arbitraire, l'algèbre de Lie $Y(n, \mathbf{m}, \Phi)$ peut être très petite, voire nulle. Considérons les deux assertions suivantes qui assurent que $Y(n, \mathbf{m}, \Phi)$ est suffisamment grosse :

(A) Pour tout $i \leq n$, il existe $\xi_i \in Y(n, \mathbf{m}, \Phi)$ tel que $\xi_i = \partial/\partial x_i + \epsilon_i$, où $\epsilon_i \in (W_{\hat{X}})_{(1)}$.

(B) $Y(n, \mathbf{m}, \Phi) \cap W(n)_{(\delta_Y)} \neq 0$, où $\delta_Y = 3$ si $Y = K$ et $\delta_Y = 2$ autrement.

La condition (A) signifie que $Y(n, \mathbf{m}, \Phi)$ engendre l'"espace tangent à l'origine".

DÉFINITION 27.— *Si l'algèbre de Lie $Y(n, \mathbf{m}, \Phi)$ satisfait aux conditions (A) et (B), alors l'algèbre de Lie $Y(n, \mathbf{m}, \Phi)^{(\infty)}$ est dite algèbre de Lie de la série de type Cartan.*

On peut maintenant expliquer la signification du symbole $\mathbf{1}$ déjà utilisé en section 3. On note $\mathbf{1}$ le n-uplet $(1, 1, \ldots)$. Il est clair que $Y(n, \mathbf{1}, Id)$ est l'algèbre de Lie notée $Y(n, \mathbf{1})$ en section 3. Comme en section 3, on peut montrer que les algèbres $Y(n, \mathbf{m}, \Phi)^{(\infty)}$ sont simples. Elles satisfont à la propriété de compatibilité :

$$Y(n, \mathbf{m}, Id)^{(2)} \subset Y(n, \mathbf{m}, \Phi)^{(\infty)} \subset Y(n, \mathbf{m}, Id).$$

7. LA CLASSIFICATION DES ALGÈBRES SIMPLES

Nous pouvons maintenant énoncer le théorème de classification, annoncé par Strade et Wilson en 1991 [**SW**], et dont la preuve complète est basée sur de nombreux articles, en particulier les 7 articles de Strade [**St1-7**].

THÉORÈME 28 (Strade et Wilson).— *Supposons $p > 7$. Alors toute algèbre de Lie simple de dimension finie est isomorphe à une algèbre classique $\mathfrak{g}_K(I)$ ou à l'une des algèbres de type Cartan $Y(n, \mathbf{m}, \Phi)^{(\infty)}$.*

Comme dans le cas restreint, les experts conjecturent que ce résultat est valable aussi en caractéristique 7. Pour $p = 5$, ils conjecturent un résultat analogue, c'est-à-dire que toute algèbre de Lie simple est classique, de type Cartan ou dans la famille de Melikian [**Me**] (cette famille n'existe qu'en caractéristique 5). À ce jour, il n' y a pas de conjecture raisonnable pour $p = 3$ et $p = 2$.

8. TECHNIQUES DE PREUVES

Dans cette section, nous n'allons dire que quelques mots sur les techniques de preuve et les difficultés de la classification.

8.1. Sous-algèbre de Cartan

Soit L une algèbre de Lie. Une sous-algèbre de Cartan est une sous-algèbre nilpotente égale à son normalisateur. Construire des sous-algèbres de Cartan est

analogue à la caractéristique zéro. On choisit un élément générique x de sorte que $ad(x)$ ait le maximum de valeurs propres non nulles. Puis on montre que la sous-algèbre de Lie

$$\mathfrak{h} = \{y \in L | ad(x)^N(y) = 0 \text{ pour } N \text{ assez grand}\}$$

est une sous-algèbre de Cartan. En caractéristique zéro, toutes les sous-algèbres de Cartan sont conjuguées (même si L n'est pas simple) et pour une algèbre simple toutes les sous-algèbres de Cartan sont ad-diagonalisables. En caractéristique finie les difficultés sont les suivantes :

(i) En général, les sous-algèbres de Cartan ne sont pas conjuguées : considérons par exemple $L = W(1,\mathbf{1}) = Der(K[x]/(x^p)$, et posons $u = 1 + x$. Les deux sous-algèbres de Cartan $\mathfrak{h}_1 = K\frac{d}{dx}$ et $\mathfrak{h}_2 = K\frac{ud}{du}$ ne sont pas conjuguées. En effet il existe beaucoup de sous-algèbres de Lie L' avec $\mathfrak{h}_1 \subset L' \subset L$, alors que $\mathfrak{h}_2$ est une sous-algèbre de Lie maximale.

(ii) En général, les sous-algèbres de Cartan ne sont pas ad-diagonalisables. Par exemple la sous-algèbre de Lie de $S(2,\mathbf{1})$ engendrée par les champs de vecteurs $x_1^j x_2^j(x_1\frac{\partial}{\partial x_1} - x_2\frac{\partial}{\partial x_2})$ pour $j = 1, 2\ldots$ n'est pas ad-diagonalisable.

(iii) Si L est restreinte, il est facile de montrer que toute sous-algèbre de Cartan est ad-triangulable. Pour des algèbres de Lie simples non restreintes, c'est un résultat très difficile qui n'est vrai que si $p \geq 7$. Pour $p = 5$, ce n'est déjà plus vrai.

8.2. Construction de grosses sous-algèbres

Ce qui différencie les algèbres de Lie classiques des algèbres de type Cartan est la présence de très grosses sous-algèbres de Lie dans les algèbres de type Cartan. Pour une algèbre de Lie L, notons $c(L)$ la codimension minimale d'une sous-algèbre de Lie. Il existe des constantes $C < C'$ telles que :

$$C(\dim L)^{1/2} \leq c(L) \leq C'(\dim L)^{1/2}, \text{ pour toute algèbre simple classique} L.$$

Par contre, lorsque L parcourt l'ensemble des algèbres simples restreintes non classiques, on a :

$$c(L) \simeq \log_p \dim L, \text{ lorsque } \dim L \to \infty.$$

Enfin il existe des algèbres de Lie L simples non restreintes de dimension arbitrairement grande avec $c(L)$ fixé : par exemple $W(1,n)$ possède une sous-algèbre de codimension 1, tandis que $\dim W(1,n) = p^n$. En résumé, les algèbres de type Cartan possèdent de très grosses sous-algèbres de Lie maximales, et elles se distinguent des algèbres classiques par cette propriété.

Par conséquent, il est naturel de chercher à construire des grosses sous-algèbres. En fait si L_0 est une sous-algèbre de codimension n, on peut plonger L dans $W(n)$, car l'espace $Hom_{L_0}(U(L), K)$ a une structure d'algèbre commutative. Notons aussi que L a une filtration naturelle qui est définie inductivement par $L_{(-1)} = L$, $L_{(0)} = L_0$ et $L_{(k)} = \{x \in L | [x, L] \subset L_{(k-1)}\}$ pour $k \geq 1$. Néanmoins pour pouvoir tirer des informations de cette construction, il est nécessaire :

1) de partir avec de grosses sous-algèbres,

2) de pouvoir déterminer quand une algèbre filtrée est une algèbre de type Cartan ou une algèbre classique.

La notion d'éléments sandwich a été développée pour construire de grosses sous-algèbres. Les réponses au second point s'appellent les théorèmes de reconnaissance. Nous allons développer ces deux points dans les alinéas suivants.

8.3. Éléments sandwich

Soit L une algèbre de Lie. Un élément $x \in L$ est dit *élément sandwich* si $ad(x)^2 = 0$. Pour des algèbres classiques, il n'y a pas d'éléments sandwich non nuls x : par le théorème de Morozov, pour tout élément nilpotent x, il existe une sous-algèbre de Lie $\mathfrak{a}$ isomorphe à $\mathfrak{sl}(2)$ qui contient x : on a donc $ad(x)^2(y) = x$ pour un certain $y \in L$.

Il est facile de vérifier que, pour tout élément sandwich y, l'exponentielle tronquée $1 + tad(x)$ est un groupe à un paramètre d'isomorphismes (ce fait requiert $p > 2$). Donc si y est un autre élément sandwich, $y + t[x, y]$ est également sandwich pour tout $t \in K$. Comme dans l'identité $ad^2(y + t[x, y]) = 0$, le terme dominant est $t^2 ad^2([x, y])$, $[x, y]$ est aussi un élément sandwich. Il s'ensuit que le sous-espace vectoriel engendré par les éléments sandwich est une sous-algèbre de Lie $S(L)$, appelée la sous-algèbre sandwich de L. Cette sous-algèbre de Lie n'est pas maximale, mais elle est un point de départ pour construire de grosses sous-algèbres.

Donnons-en un exemple. Partons de l'algèbre de Lie $L = W(1, 1)$ et posons $L_i = x^{i+1} \frac{d}{dx}$ pour $-1 \leq i \leq p - 2$. On a $[L_i, L_j] = (j - i)L_{i+j}$ si $i + j \leq p - 2$ et $[L_i, L_j] = 0$ sinon. L'algèbre $S(L)$ a pour base $(L_i)_{i \geq \frac{p+1}{2}}$. Le normalisateur de $S(L)$ est donc l'algèbre de Lie de codimension 1 engendrée par tous les L_i avec $i \geq 0$. On voit donc dans ce cas qu'on a un moyen naturel pour construire une grosse sous-algèbre.

8.4. Théorèmes de reconnaissance

Les algèbres classiques ont été caractérisées par Mills et Seligman. Soit L une algèbre de Lie contenant une sous-algèbre de Cartan H ad-diagonalisable. Pour tout

$\alpha \in H^* \setminus 0$ posons $L_\alpha = \{x \in L | ad(h)(x) = \alpha(h)x \text{ pour tout } h \in H\}$ et soit $\Delta = \{\alpha \in H^* \setminus 0 | L_\alpha \neq 0\}$. Par conséquent L se décompose en $L = H \oplus \oplus_{\alpha \in \Delta} L_\alpha$. On rappelle l'hypothèse $p > 3$.

THÉORÈME 29 (Mills et Seligman [**MS**]).— *Supposons que :*

(i) $Z(L) = 0$ *et* $L^{(1)} = L$;

(ii) *pour tout* $\alpha \in \Delta$, $[L_\alpha, L_{-\alpha}]$ *est de dimension* 1 ;

(iii) *pour tout* $\alpha, \beta \in \Delta$, *il existe* $k \in \mathbf{F_p}$ *tel que* $\alpha + k\beta \notin \Delta \cup \{0\}$.

Alors L *est une somme directe d'algèbre de Lie simple classique.*

Les hypothèses (i) et (ii) sont naturelles. Expliquons brièvement à quoi sert l'hypothèse (iii). Pour $\alpha \in \Delta$, posons $H_\alpha = [L_\alpha, L_{-\alpha}]$. En fait le début de la preuve consiste à montrer que la sous-algèbre de Lie $L_\alpha \oplus H_\alpha \oplus L_{-\alpha}$ est isomorphe à $\mathfrak{sl}(2)$.

Le premier pas consiste à prouver que $\alpha(H_\alpha) \neq 0$. Si ce n'est pas le cas, on choisit $x \in L_\alpha$, $y \in L_{-\alpha}$ avec $[x, y] \neq 0$. Posons $h = [x, y]$. On a donc $\alpha(h) \neq 0$. Comme le centre de L est non nul, il existe $\beta \in \Delta$ avec $\beta(h) \neq 0$. Grâce à l'hypothèse (iii), on peut en outre supposer que $\beta + \alpha \notin \Delta$. Soit $v \in L_\beta$ un élément non nul. Comme $[x, v]$ appartient à $L_{\alpha+\beta} = 0$, on a $[x, v] = 0$. Pour tout $i \leq p - 1$, on pose $v_i = ad(y)^i(v)$. Il est clair que v_i appartient à $L_{\alpha+i\beta}$ et un calcul facile montre que $[x, v_i] = i\beta(h)v_{i-1}$, donc aucun des v_i n'est nul. Or v_{p-1} appartient à $L_{\beta+\alpha}$, ce qui contredit l'hypothèse (iii).

On peut donc choisir $x \in L_\alpha$, $y \in L_{-\alpha}$ avec $[x, y] \neq 0$, et l'algèbre de Lie engendrée par x et y est isomorphe à $\mathfrak{sl}(2)$. Par des calculs analogues, on établit que $L_{\pm\alpha}$ est de dimension 1, et aussi que $L_{n\alpha} = 0$ pour $n \neq -1, 0$, ou 1.

Les théorèmes de reconnaissance des algèbres de type Cartan sont plus complexes. En fait ils utilisent une technique de filtration plus compliquée que celle décrite plus haut. Partons d'une algèbre de Lie L, d'une sous-algèbre $L_{(0)}$ et d'un $L_{(0)}$-sous-module $L_{(-1)} \supset L_{(0)}$ tel que $L_{(-1)}$ engendre L (auparavant, on n'avait considéré que le cas $L_{(1)} = L$). Cette condition est automatiquement satisfaite si $L_{(0)}$ est maximale. On construit une filtration $L_{(s)}$ de L comme suit : pour définir la partie négative de la filtration, on pose :

$$L_{(-2)} = L_{(-1)} + [L_{(-1)}, L_{(1)}],$$
$$L_{(-3)} = L_{(-2)} + [L_{(-1)}, L_{(-2)}], \ \ldots,$$
$$L_{(-(s+1))} = L_{(-s)} + [L_{(-1)}, L_{(-s)}].$$

Pour définir la partie positive de la filtration, on pose : $L_{(1)} = \{x \in L_{(0)} | [x, L_{(-1)}] \subset L_{(0)}\}$, $L_{(2)} = \{x \in L_{(1)} | [x, L_{(-1)}] \subset L_{(1)}\}$, ..., $L_{(s+1)} = \{x \in L_{(s)} | [x, L_{(-1)}] \subset L_{(s)}\}$. Cette filtration est appelée la *filtration standard* associée à la paire $(L_{(0)}, L_{(-1)})$.

THÉORÈME 30 (Kac **[Kc2]**, Wilson **[W1]**, Benkart, Gregory, Premets) ($p > 3$).— *Soit L une algèbre de Lie simple, et soit $(L_{(0)}, L_{(-1)})$ une paire telle que $L_{(0)}$ soit une sous-algèbre maximale et $L_{(-1)}/L_{(0)}$ soit un $L_{(0)}$-module simple. Soient s, s' les plus petits entiers tels que la filtration standard soit de la forme : $L = L_{(-s')} \supset \ldots \supset L_{(s)} \supset L_{(s+1)} = 0$. Posons $G =^{gr} L$, et supposons :*

((i) *G_0 est une somme directe d'idéaux où chacun d'entre eux est une algèbre simple classique, abélienne ou isomorphe à $\mathfrak{gl}(pn)$, $\mathfrak{pgl}(pn)$, ou à $\mathfrak{sl}(pn)$.*

(ii) *Pour tout $j \geq 0$, et tout $j \in G_{-j} \setminus 0$, $[x, G_1] \neq 0$,*

(iii) *$s' \leq s$.*

Alors L est une algèbre de Lie classique, une algèbre de type Cartan, ou une algèbre de Melikian si $p = 5$.

Pour des questions de taille de cet exposé écrit, nous n'avons pas décrit les techniques les plus sophistiquées développées par Strade et Wilson ces dernières années, comme la classification des 2-sections, le théorème de structure des 3-sections et les techniques de tores maximaux dans les p-enveloppes. On pourra se reporter à **[St7]** pour un exposé de survey.

BIBLIOGRAPHIE

[BGOSW] G.M. Benkart, T.B. Gregory, J.M. Osborn, H. Strade et R.L. Wilson - *Isomorphism classes of Hamiltonian Lie algebras.* In *Lie algebras*, Madison 1987, Springer. Lecture Notes in Math. **1373** (1989), 42-57.

[BOS] G.M. Benkart, J.M. Osborn et H. Strade - *Contributions to the classification of simple modular Lie algebras*, Trans. Amer. Math. Soc. **341** (1994), 227-252.

[BW1] R.E. Block et R.L. Wilson - *The restricted simple Lie algebras are of classical or Cartan type*, Proc. Nat. Acad. Sci. U.S.A. **81** (1984), 5271-5274.

[BW2] R.E. Block et R.L. Wilson - *Restricted simple Lie algebras.* In *Lie algebras and related topics* (Windsor, Ont., 1984), Amer. Math. Soc. CMS Conf. Proc. **5** (1986), 3-17.

[BW3] R.E. Block et R.L. Wilson - *Classification of the restricted simple Lie algebras*, J. Algebra **114** (1988), 115-259.

[Cn] E. Cartan, *Les groupes de transformations continus, infinis, simples*, Ann. Sci. Ecole Norm. Sup. **26** (1909), 93-161.

[Cr] P. Cartier - *Une nouvelle opération sur les formes différentielles*, C. R. Acad. Sci. Paris **244** (1957), 426-428.

[J1] N. Jacobson - *Abstract derivation and Lie algebra*, Trans. Amer. Math. Soc. **42** (1937), 206-224.

[J2] N. Jacobson - *Lie algebras*, Interscience Tracts in Pure and Applied Mathematics, **10** (1962).

[GM] I. M. Gelfand et O. Mathieu - *On the cohomology of the Lie algebra of hamiltonian vector fields*, J. of Funct. Anal. **108** (1992), 347-360.

[G] V. Guillemin - *Infinite dimensional primitive Lie algebra*, J. differential geometry **4** (1970), 257-282.

[Kc1] V. G. Kac - *Simple graded Lie algebras of finite growth*, Math. USSR Izv. **2** (1968), 1271-1311.

[Kc2] V. G. Kac - *The classification of the simple Lie algebras over a field with non-zero characteristic*, Izv. Akad. Nauk SSSR Ser. Mat. **34** (1970), 385-408.

[Kc3] V. G. Kac - *Global Cartan pseudogroups and simple Lie algebras in characteristic p* (en russe), Usp. Math. Nauk 26 (1971), 199-200.

[Kz] N. M. Katz - *A conjecture in the arithmetic theory of differential equations*, Bull. Soc. Math. France **110** (1982), 203-239.

[KS] A. I. Kostrikin et I. R. Shafarevitch - *Cartan pseudo-groups and Lie p-algebras*, Soviet Dokl. **7** (1966), 715-718.

[Ma] O. Mathieu - *Classification of simple graded Lie algebras of finite growth*, Inv. Math. **108** (1992), 455-519.

[Me] G. M. Melikian -*Simple Lie algebras of characteristic* 5, Uspekhi Mat. Nauk **35** (1980), 203-204.

[MS] W. H. Mills et G.B. Seligman -*Lie algebras of classical type*, J. Math. Mech. **6** (1957), 519-548.

[PS] A. Premet et H. Strade - *Simple Lie algebras of small characteristic*, I, Sandwich elements. J. Algebra **189** (1997), 419-480.

[Se1] G.B. Seligman - *On a class of semisimple restricted Lie algebras*, Proc. Nat. Acad. Sci. U.S.A. **40** (1954), 726-728.

[Se2] G.B. Seligman - *A survey of Lie algebras of characteristic p*, Report of a conference on linear algebras, National Academy of Sciences-National Research Council. **502** (1956), 24-32.

[Se3] G.B. Seligman - *Some remarks on classical Lie algebras*, J. Math. Mech. **6** (1957), 549-558.

[Stb] R. Steinberg - *Lectures on Chevalley groups.* Notes, Yale University (1967).

[St1] H. Strade - *The classification of the simple modular Lie algebras*, III. Solution of the classical case. Ann. of Math. **133** (1991), 577-604.

[St2] H. Strade - *The classification of the simple modular Lie algebras: a revised approach, Hadronic mechanics and nonpotential interactions*, Part 1 (Cedar Falls, IA, 1990), Nova Sci. Publ. (1992), 79-99.

[St3] H. Strade - *The classification of the simple modular Lie algebras*, II. The toral structure. J. Algebra **151** (1992), 425-475.

[St4] H. Strade - *The classification of the simple Lie algebras over algebraically closed fields with positive characteristic, Jordan algebras* (Oberwolfach 1992), de Gruyter (1994), 281-299.

[St5] H. Strade - *The classification of the simple modular Lie algebras, V. Algebras with Hamiltonian two-sections*, Abh. Math. Sem. Univ. Hamburg **64** (1994), 167-202.

[St6] H. Strade - *The classification of the simple modular Lie algebras. VI. Solving the final case*, Trans. Amer. Math. Soc. **350** (1998), 2553-2628.

[St7] H. Strade - *The classification of simple modular Lie algebras*, Preprint.

[SW] H. Strade et R. L. Wilson - *Classification of simple Lie algebras over algebraically closed fields of prime characteristic*, Bull. Amer. Math. Soc. **24** (1991), 357-362.

[T1] J. Tits - Résumé de cours, annuaire du Collège de France (1980-1981), 75-86, et (1981-1982), 91-105.

[T2] J.Tits - *Groupes associés aux algèbres de Kac-Moody*, Sém. Bourbaki, Exp. n° 700 (novembre 1988), Astérisque **177-178** (1989), 7-31.

[T3] J. Tits - *Group and group functors attached to Kac-Moody data*, Lecture Notes in Math. **1111**, Springer-Verlag (1985), 193-223.

[W1] R. L. Wilson - *A structural characterization of the simple Lie algebras of generalized Cartan type over fields of prime characteristic*, J. Algebra **40** (1976), 418-465.

[W2] R. L. Wilson - *Simple Lie algebras of type S*, J. Algebra **62** (1980), 292-298.

[W3] R. L. Wilson - *The classification problem for simple Lie algebras of characteristic p*. In *Algebra*, Proc. Conf. Carbondale 1980, Springer. Lecture Notes in Math. **848** (1981), 13-32.

Olivier MATHIEU

Université Louis Pasteur
IRMA
7, rue rené Descartes
67084 Strasbourg Cedex
mathieu@math.u-strasbg.fr

Université Paris 7
U.F.R. de mathématiques
2, place Jussieu
75251 Paris Cedex 05

Astérisque

MICHEL TALAGRAND

Verres de Spin et optimisation combinatoire

Astérisque, tome 266 (2000), Séminaire Bourbaki,
exp. n° 859, p. 287-317

<http://www.numdam.org/item?id=SB_1998-1999__41__287_0>

© Société mathématique de France, 2000, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (http://smf4.emath.fr/Publications/Asterisque/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Séminaire BOURBAKI Mars 1999
51ème année, 1998-99, n° 859

VERRES DE SPIN ET OPTIMISATION COMBINATOIRE

par **Michel TALAGRAND**

1. INTRODUCTION

Les verres de spin sont un vaste sujet, touchant aux probabilités, à l'optimisation combinatoire, et bien sûr à la physique. Les physiciens ont sans doute écrit plus de mille pages sur ce sujet, à commencer par le livre de M. Mézard, G. Parisi, M. Virasoro, "Spin Glass Theory and Beyond". Ce livre est remarquable non seulement par le foisonnement des idées qu'il introduit, mais peut-être, pour nous mathématiciens il l'est plus encore par son sujet. Il présente en effet l'étude d'objets motivés par la physique, mais purement mathématiques, et qui sont souvent, comme nous allons le voir, de nature extrêmement simple et canonique. Cette étude repose sur des méthodes de physique théorique, et est de plus exposée dans le langage de cette discipline, ce qui n'en facilite pas particulièrement l'accès pour un mathématicien. Les mathématiciens ont maintenant écrit sans doute plus de cinq cents pages sur le sujet. Il est très frappant toutefois de constater que les questions traitées à grand peine dans ces travaux n'occupent pas plus de quelques pages (d'introduction...) des publications des physiciens.

Le présent texte se voudrait une introduction accessible au sujet. Pour cela, il contiendra peu d'énoncés techniques, voire même peu d'énoncés tout court. Il tentera d'expliquer la nature des objets dont on parle, et quelques problèmes naturels. Il tentera aussi de donner un aperçu de quelques méthodes, des difficultés auxquelles elles font face, et d'expliquer le contexte et la signification des progrès récents sur les mystérieuses formules dites "réplique-symétrique". Il tentera aussi bien sûr d'expliquer la situation hautement paradoxale décrite plus haut. Pour un exposé plus détaillé, on pourra consulter un autre texte d'introduction plus ambitieux [T5], disponible en particulier sur la page personnelle de l'auteur [T6]. On y trouvera des énoncés précis et des indications plus complètes sur certains schémas de preuves.

2. QUELQUES PROBLÈMES FACILES (À POSER)

Le cube discret $\Sigma_N = \{-1,1\}^N \subset \mathbb{R}^N$ sera un objet central de cet exposé.

Étant donné un sous-espace A de $\mathbb{R}^N$, dénotons par Q_A la projection orthogonale sur A, et par $\|.\|$ la norme euclidienne. Considérons la quantité

$$R_A = \frac{1}{\sqrt{N}} \max_{\boldsymbol{\sigma} \in \Sigma_N} \|Q_A(\boldsymbol{\sigma})\| .$$

On observe que $\|\boldsymbol{\sigma}\| = \sqrt{N}$, ce qui justifie le facteur de normalisation $N^{-1/2}$.

Fixons maintenant $M \leq N$. L'ensemble des sous-espaces de $\mathbb{R}^N$ de dimension M est muni d'une probabilité naturelle ; désignons par E_M l'espérance vis-à-vis de cette probabilité. La quantité $Q(N,M) = E_M R_A$ est donc un nombre dépendant seulement de M et N. Dénotant par $\lfloor x \rfloor$ la partie entière de x, étant donné $0 < \alpha < 1$, on a la question naturelle suivante.

Problème 2.1.— Est-il vrai que la limite

$$\lim_{N \to \infty} Q(N, \lfloor \alpha N \rfloor)$$

existe ?

Cette question est étroitement reliée, tant au niveau philosophique qu'au niveau mathématique, aux problèmes qui seront considérés ultérieurement.

Au niveau philosophique, tout le monde s'accordera sans doute à penser que ne voyant aucune raison pour que cette limite n'existe pas, il serait vraiment surprenant que cela ne soit pas le cas.

On peut imaginer par contre que la plupart des mathématiciens et la plupart des physiciens ne donneront pas la même réponse à la question plus pratique de savoir si l'on doit travailler pour essayer d'éliminer tout à fait la lointaine possibilité que cette limite n'existe pas.

Au niveau mathématique, étant donné $\boldsymbol{\sigma} \in \Sigma_N$, on peut considérer la variable aléatoire $X_{\boldsymbol{\sigma}}$ (définie sur l'ensemble des sous-espaces de $\mathbb{R}^N$ de dimension N muni de sa probabilité canonique) donnée par $X_{\boldsymbol{\sigma}}(A) = \|Q_A(\boldsymbol{\sigma})\|$. Le problème serait facile si la famille des variables aléatoires $X_{\boldsymbol{\sigma}}(\boldsymbol{\sigma} \in \Sigma_N)$ était indépendante. Cela n'est pas le cas.

Toutefois, pour des raisons de symétrie, la corrélation de $X_{\boldsymbol{\sigma}}$ et $X_{\boldsymbol{\sigma}'}$ ne dépend que de la distance de Hamming de $\boldsymbol{\sigma}$ et $\boldsymbol{\sigma}'$ (c'est-à-dire du nombre de coordonnées où $\boldsymbol{\sigma}$ et $\boldsymbol{\sigma}'$ diffèrent). On est donc en droit d'espérer que, pour $N \to \infty$, un ordre se dégage de cette structure de corrélation assez régulière, ce qui bien sûr renforce

l'espoir d'une réponse positive au problème 2.1. Cette idée d'un ordre émergeant d'une structure de corrélation régulière est sans doute l'un des thèmes centraux du sujet.

Examinons maintenant des questions similaires, plus proches des verres de spins, où les projections orthogonales sont remplacées par des opérateurs définis par des matrices aléatoires. Considérons l'opérateur aléatoire T de $\mathbb{R}^M$ dans $\mathbb{R}^N$, de matrice $(g_i^k)_{i\leq N, k\leq M}$, où les g_i^k sont des variables indépendantes normales standard. La norme de T est bien comprise si $\mathbb{R}^M$ et $\mathbb{R}^N$ sont munis de la structure euclidienne. Mais il n'en est pas de même si $\mathbb{R}^M$ est muni de la norme euclidienne et $\mathbb{R}^N$ de sa norme ℓ^1. La norme d'opérateur $R(M,N)$ de T (beaucoup plus difficile à calculer dans ce nouveau cas) est alors

$$R(M,N) = \sup\left\{ \sum_{i\leq N} \Big| \sum_{k\leq M} g_i^k x_k \Big| ; \sum_{k\leq M} x_k^2 = 1 \right\} ; \tag{2.1}$$

ou, par dualité

$$R(M,N) = \sup\left\{ \left(\sum_{k\leq M} \Big| \sum_{i\leq N} g_i^k \sigma_i \Big|^2 \right)^{1/2} ; \boldsymbol{\sigma} \in \Sigma_N \right\}. \tag{2.2}$$

On remarque que ces deux quantités s'expriment comme des sup de variables aléatoires indexées par des ensembles ayant une structure géométrique simple.

Problème 2.2.— Est-il vrai que

$$\lim_{N\to\infty} \frac{1}{N}\, ER(\lfloor \alpha N \rfloor, N)$$

existe pour tout $\alpha > 0$?

Ici E désigne bien sûr l'espérance dans les variables (g_i^k).

3. LE MODÈLE SK À TEMPÉRATURE ZÉRO

Ce paragraphe mime l'introduction de [M-P-V].

Considérons une vaste population d'individus, numérotés de 1 à N. Ils se connaissent tous. Les sentiments respectifs des individus $i < j$ sont mesurés par un nombre g_{ij}. (On suppose donc la symétrie : les sentiments de i pour j et de j pour i sont égaux.) On veut modéliser une situation où ces sentiments respectifs sont aléatoires.

On suppose (ce qui représente bien entendu une simplification extrême par rapport à tout modèle réaliste !) que la famille de variables aléatoires (g_{ij}) est indépendante. Puisque l'on cherche à écrire un modèle simple, on supposera que les variables g_{ij} sont gaussiennes standard. Ces variables aléatoires, qui vont jouer un rôle central, seront appelées le désordre, car c'est cela qu'elles modèlent. À cause de l'indépendance, pour beaucoup de triplets i, j, k, i est ami de j et de k $(g_{ij}, g_{ik} > 0)$ mais j et k sont ennemis $(g_{ik} < 0)$, une situation connue sous le nom de frustration. Cela crée bien sûr des tensions. Pour les soulager, essayons de séparer la population en deux groupes, regroupant les amis et séparant les ennemis autant que faire se peut. Pour cela, il est pratique d'assigner à chaque individu i un nombre $\sigma_i \in \{-1, 1\}$. Chaque configuration $\boldsymbol{\sigma}$ de Σ_N définit ainsi naturellement une partition de la population en deux classes. Pour mesurer l'efficacité de cette partition à regrouper les amis et à séparer les ennemis, il est naturel de considérer la quantité

$$A_N(\boldsymbol{\sigma}) = \sum_{1 \le i < j \le N} g_{ij}\sigma_i\sigma_j \tag{3.1}$$

qui ajoute les interactions de chaque paire d'individus dans le même groupe, mais retranche les interactions des individus dans des groupes différents.

Nous nous intéressons donc à $\max_{\boldsymbol{\sigma}} A_N(\boldsymbol{\sigma})$ et en particulier à son espérance (par rapport au désordre bien sûr). Les variables aléatoires $A_N(\boldsymbol{\sigma})$ sont conjointement gaussiennes, c'est-à-dire forment un processus gaussien. Le problème de la mesure de la "taille" d'un processus gaussien par l'estimation de l'espérance de son sup est un problème important qui a été beaucoup étudié. On sait (au moins en principe) estimer cette quantité en fonction de la covariance du processus [F], [T1]. Ces estimations ne sont hélas précises qu'à une constante multiplicative près, et dans le cas présent nous apprennent seulement que $E\max_{\boldsymbol{\sigma}} A_N(\boldsymbol{\sigma})$ est d'ordre $N^{3/2}$ (ce qui n'est d'ailleurs pas bien difficile). On a donc le problème naturel suivant

Problème 3.1.— Est-il vrai que la limite

$$\lim_{N\to\infty} N^{-3/2} E \max_{\boldsymbol{\sigma}} A_N(\boldsymbol{\sigma})$$

existe ?

A en croire les simulations numériques (portant sur des valeurs de N de l'ordre de 100) c'est bien le cas et cette limite vaut environ 0,7633.

Il est tout à fait intéressant de calculer la covariance du processus (3.1). On a

$$EA_N(\boldsymbol{\sigma})A_N(\boldsymbol{\sigma}') = \sum_{i<j} \sigma_i\sigma_j\sigma'_i\sigma'_j = \frac{1}{2}\left(\sum_{i\le N} \sigma_i\sigma'_i\right)^2 - N, \tag{3.2}$$

une dépendance algébrique simple en $\boldsymbol{\sigma} \cdot \boldsymbol{\sigma}' = \sum_{i \leq N} \sigma_i \sigma_i'$ ce qui bien sûr renforce l'espoir de l'émergence d'un ordre lorsque $N \to \infty$.

Le problème 3.1 est très proche en esprit du problème 2.1. En effet, si l'on désigne par $(\lambda_k)_{k \leq N}$ et $(\boldsymbol{x}_k)_{k \leq N}$ respectivement les valeurs propres et les vecteurs propres de la matrice symétrique (a_{ij}) donnée par $a_{ij} = \frac{1}{2} g_{ij}$ si $i < j, a_{ij} = \frac{1}{2} g_{ji}$, si $i > j$, et $a_{ii} = 0$, alors

$$\sum_{i<j} g_{ij} \sigma_i \sigma_j = \sum_{i,j} a_{ij} \sigma_i \sigma_j = \sum_{k \leq N} \lambda_k (\boldsymbol{x}_k \cdot \boldsymbol{\sigma})^2$$

et il nous faut donc comprendre $E \max_{\boldsymbol{\sigma}} \sum_{k \leq N} \lambda_k (\boldsymbol{x}_k \cdot \boldsymbol{\sigma})^2$.

La transformation orthogonale aléatoire qui envoie la base canonique $\mathbb{R}^N$ sur les $(\boldsymbol{x}_k)_{k \leq N}$ est indépendante des (λ_k) et uniformément distribuée sur le groupe orthogonal, comme il résulte de l'invariance de la distribution de la matrice (a_{ij}) par changement de base orthonormale. D'autre part (d'après la célèbre "loi du demi cercle") les nombres (λ_k) dépendent peu des (g_{ij}). Le problème 2.1 correspond tout simplement au cas où les λ_k, au lieu, comme ici, de suivre la loi du demi-cercle, ne prendraient que les valeurs 0 et 1, $\lfloor \alpha N \rfloor$ d'entre elles prenant la valeur 1.

4. OPTIMISATION COMBINATOIRE ET MÉCANIQUE STATISTIQUE

Il convient de remplacer la quantité (2.1) par la quantité plus appropriée

$$H_N(\boldsymbol{\sigma}) = -\frac{1}{\sqrt{N}} \sum_{1 \leq i < j \leq N} g_{ij} \sigma_i \sigma_j. \tag{4.1}$$

Le signe moins respecte les conventions de la physique, pour laquelle $H_N(\boldsymbol{\sigma})$ représente l'énergie de la configuration $\boldsymbol{\sigma}$, et qui cherche toujours à minimiser, et non pas maximiser l'énergie. Le facteur $N^{-1/2}$ introduit la normalisation nécessaire pour que $\min_{\boldsymbol{\sigma}} H_N(\boldsymbol{\sigma})$ soit d'ordre N, ce qui sera utile pour la suite. La section 2 s'intéressait au calcul de $\min_{\boldsymbol{\sigma}} H_N(\boldsymbol{\sigma})$, c'est-à-dire (par définition) en termes physiques, à l'énergie de l'état fondamental.

La mécanique statistique propose de remplacer le problème de trouver la plus petite valeur de $H_N(\boldsymbol{\sigma})$ (ce que l'on appelle d'ordinaire un problème d'optimisation combinatoire) par l'étude "des plus petites valeurs" de $H_N(\boldsymbol{\sigma})$, avec une pondération adéquate. Cette pondération utilise un paramètre $\beta \geq 0$ et introduit la mesure de Gibbs G_N, une mesure de probabilité qui donne à $\boldsymbol{\sigma}$ le poids

$$G_N(\boldsymbol{\sigma}) = \frac{1}{Z_N} \exp(-\beta H_N(\boldsymbol{\sigma})), \tag{4.2}$$

où Z_N est le facteur de normalisation $\sum_{\boldsymbol{\sigma}} \exp(-\beta H_N(\boldsymbol{\sigma}))$. Si l'on fixe N, pour $\beta \to \infty$, la mesure de Gibbs se concentre sur l'état fondamental. On peut donc espérer obtenir de l'information sur l'énergie de cet état si l'on arrive à étudier la mesure de Gibbs pour β grand.

La difficulté de l'étude de la mesure de Gibbs provient du facteur Z_N dans (4.2). Celui-ci est la somme de 2^N termes d'ordre de grandeurs très différents. Beaucoup de ces termes sont ≤ 1 (si $H_N(\boldsymbol{\sigma}) \geq 0...$) ; les plus grands d'entre eux sont d'ordre exponentiel, puisque $\max(-H_N(\boldsymbol{\sigma}))$ est d'ordre N. Il y a peu de grands termes, mais de nombreux petits. La normalisation est choisie de sorte qu'il ne soit nullement évident de quel ordre sont les termes qui contribuent le plus : cela dépend d'ailleurs de la valeur de β.

Nous avons présenté ici la mécanique statistique comme approche à l'optimisation combinatoire ; elle va en effet nous permettre d'aborder nombre de problèmes purement mathématiques faisant intervenir du "désordre" mais n'ayant a priori rien à voir avec l'étude de la matière. Il nous semble que la possibilité d'approcher ces problèmes contribue de façon essentielle à l'attrait du sujet. Mais bien entendu la mécanique statistique présente un intérêt considérable par elle-même ! Si un système physique peut se trouver dans l'une des configurations $\boldsymbol{\sigma} \in \Sigma_N$, et qu'il possède une énergie $H_N(\boldsymbol{\sigma})$ dans cette configuration, la mesure de Gibbs (4.2) donne la probabilité que le système en équilibre thermique à température $1/\beta$ se trouve dans la configuration $\boldsymbol{\sigma}$. C'est évidemment la mécanique statistique qui a conduit Sherrington et Kirkpatrick à introduire leur célèbre modèle. Le but de celui-ci est de reproduire au moins le comportement qualificatif de certains alliages qui présentent d'étranges propriétés magnétiques à basse température, les célèbres "verres de spin". Dans un tel alliage, par exemple 5% Fe et 95% Au, seuls les atomes de fer ont des propriétés magnétiques. Du fait de leur grande dilution, la distance d'un atome de fer et de ses plus proches voisins est assez aléatoire. Il en est donc de même de leurs interactions, et l'on pense que c'est le caractère aléatoire de ses interactions qui est la source des comportements étranges observés expérimentalement. Sherrington et Kirkpatrick conservent le caractère aléatoire des interactions, le simplifiant à l'extrême en supposant l'indépendance des interactions pour des paires distinctes. Le modèle ne tient aucun compte de la position des atomes dans l'espace, puisqu'il n'y a pas de notion de "proximité" parmi les individus. C'est là l'approximation du "champ moyen" courante en physique. (Notons en passant que si les modèles à champ moyen sont souvent faciles à étudier, ce n'est pas du tout le cas des modèles désordonnés rencontrés dans cet exposé.) Il existe d'autres modèles pour les verres de spins, qui cette fois tiennent compte des positions

géométriques des individus, que l'on suppose localisés sur un réseau régulier, et où seuls les proches voisins interagissent. Ces modèles, comme on doit s'y attendre, sont encore beaucoup plus difficiles à étudier, à tel point qu'à basse température il ne s'est pas encore dégagé de consensus sur leur comportement probable, même au niveau des simulations numériques. Il est bien sûr extrêmement difficile de dire quoi que ce soit au niveau mathématique (voir cependant [N-S]). Nous ne parlerons pas davantage de ces modèles, qui constituent d'ailleurs un autre sujet.

5. TRAVAUX DES PHYSICIENS

Nous avons expliqué pourquoi ceux-ci se sont intéressés au modèle SK. Celui-ci semblait facile, comme l'indique le titre "a solvable model..." de l'article original. La situation n'était pas aussi simple. Après plusieurs essais une solution a été proposée par G. Parisi. Ses premières présentations faisaient appel à un formalisme, dont la signification demeure particulièrement obscure, dit "méthode des répliques". Il a fallu de nombreux travaux pour éclaircir le sens de ces idées. La solution proposée par Parisi se révèle d'une grande subtilité et est devenue à juste titre célèbre. Les nouveaux concepts auxquels elle mène ont été appliqués avec grand succès à nombre de problèmes, comme cela est décrit avec brio dans [M-P-V].

Du point de vue des physiciens, il semble n'y avoir plus guère de doute que les concepts proposés par Parisi représentent bien le comportement réel du modèle SK, et les plus brillants d'entre eux sont depuis longtemps allés explorer de nouveaux sujets. Les très nombreux travaux encore publiés dans le domaine concernent principalement des variations des principaux modèles, qui sont le plus souvent "résolus" en calculant par optimisation les différents paramètres des solutions proposées par Parisi. Telle est en gros la situation pour les problèmes "à basse température", c'est-à-dire les problèmes difficiles. En ce qui concerne les problèmes "à haute température" (nous donnerons un sens à cela plus tard), on se contente au mieux d'effectuer un rapide calcul par le formalisme des répliques, dans sa version la plus simple dite "réplique-symétrique". Or, en dehors de quelques situations très particulières, ce n'est que tout récemment que ce dernier cas a pu être abordé mathématiquement.

En ce qui concerne les différences de méthodes entre mathématiciens et physiciens, il serait sans doute souhaitable de pouvoir s'en tenir au commentaire suivant le problème 2.1. Cette question a toutefois une telle importance dans ce sujet qu'il semble préférable à l'auteur de discuter ce point avec beaucoup de soin quitte à risquer, sinon son existence, du moins d'être taxé de lourdeur. Tâchons donc d'expliquer la na-

ture de l'approche utilisée par les physiciens dans ce domaine. Ceux-ci n'exigent pas des arguments irréfutables, mais cherchent plutôt (ce qui est beaucoup plus sensé) à comprendre le comportement des objets qu'ils étudient avec un grand degré de certitude. Cette démarche les amène donc à supposer toujours a priori que "tout se passe bien" et à ne pas prêter attention aux pathologies concevables mais peu vraisemblables. Par exemple dans [M-P-V] on rencontre souvent l'argument de "limite thermodynamique" qui consiste essentiellement à supposer que toutes les quantités rencontrées ont des limites quand $N \to \infty$, et qui bien sûr résout facilement les problèmes 2.1 et 2.2. Les physiciens proposent toujours comme solution le comportement le plus naturel. "Le plus naturel" ne signifie absolument pas que cela est facile. Il a évidemment fallu une ingéniosité considérable pour découvrir ce qui nous est proposé dans [M-P-V]. Une fois ce comportement naturel découvert il est soumis à une analyse de cohérence interne heuristique, mais très serrée, puis il est confronté aux simulations numériques. Même si celles-ci sont difficiles à effectuer et ne peuvent pas avoir lieu pour N grand, elles s'efforcent toujours de tester des comportements dont la coïncidence avec les prédictions théoriques ne saurait être le fait du hasard. Lorsque plus personne n'a d'objections à formuler, le consensus est établi que la solution proposée reflète bien la réalité. Les physiciens semblent n'avoir aucun état d'âme à effectuer des calculs qui semblent assez formels, tels que des développements limités sans contrôle des termes d'erreur. Tout cela est justifié a posteriori par ce qui compte vraiment : l'adéquation des prévisions aux simulations numériques.

Les physiciens font de plus appel à une arme redoutable : une intuition foudroyante, qui s'appuie littéralement sur des siècles d'expérience collective. Le piment de la situation est que cette intuition est ici appliquée à des objets mathématiques tels le modèle SK, dont on ne peut pas dire qu'il soit un modèle réaliste de comportement de la matière, et qui est donc assez loin des exemples sur lesquels cette intuition s'est construite. L'une de ces intuitions est qu'à "haute température, un système est toujours dans un état pur" (expression dont le sens sera expliqué plus loin). Cela va de soi pour eux (au même titre que la solution au problème 2.1), et ils pensent donc qu'il s'agit d'une situation très simple. L'ironie du sort a voulu que du point de vue mathématique il semble extrêmement difficile de justifier cette situation "d'état pur".

Afin de mettre en lumière le fait que les physiciens ne font pas des mathématiques (ce qui évidemment n'est pas un reproche !) il est tentant de mettre en exergue la méthode dite des répliques, un formalisme mystérieux faisant intervenir des entiers $n, 0 < n < 1$, des matrices de dimension 0×0, des fonctions d'un nombre négatif de variables, etc. Il faut souligner toutefois que les succès de cette méthode (qui est tout

simplement une forme d'interpolation particulièrement audacieuse) ont semblé tout aussi mystérieux aux physiciens qu'ils ne le semblent à nous-même, et ceux-ci ont, dans plusieurs cas importants "vérifié" les prédictions de cette méthode par une approche plus compréhensible (la méthode dite de la cavité, qui n'est toutefois pas utilisée de façon rigoureuse comme nous le ferons plus loin). Ils se sont ainsi convaincus que la méthode des répliques fournit les bonnes réponses. Comme celle-ci est très puissante, c'est-à-dire propose par des calculs assez simples des expressions compliquées que l'on aurait plus de difficultés à obtenir autrement (même au niveau non rigoureux), ils ne voient donc bien sûr aucune raison de ne pas l'utiliser systématiquement.

6. LE MODÈLE SK : PREMIERS RÉSULTATS

Le modèle SK *n'est pas* notre centre d'intérêt principal. (De quoi est-ce un modèle ?) Il y a toutefois plusieurs raisons d'en faire le fil directeur du début de notre exposé : il est très connu, il est historiquement important, et il est bien plus simple que les autres modèles que nous considérerons plus tard.

Quelles questions faut-il poser ? La caractéristique principale du modèle est naturellement que les interactions, et donc l'hamiltonien H_N sont aléatoires. On doit bien sûr s'intéresser à la structure de G_N pour les réalisations typiques du désordre. Puisque la fonction de partition Z_N est difficile à estimer, on doit aussi s'intéresser à sa valeur. Pour obtenir une quantité d'ordre 1, il est naturel d'étudier plutôt $N^{-1}\log Z_N$. Ceci est une quantité aléatoire. Par bonheur, des principes généraux, et plus particulièrement le phénomène dit "de concentration de la mesure", font que les fluctuations de $N^{-1}\log Z_N$ dues au désordre sont l'ordre au plus $N^{-1/2}$ (au sens très fort d'une inégalité exponentielle) et donc l'essentiel de l'information sur Z_N est contenu dans le nombre

$$p_N = p_N(\beta) = E\frac{1}{N}\log Z_N. \tag{6.1}$$

Dans [A-L-R], on appelle cette quantité "l'énergie libre" (moyenne par site) du système, convention suivie ensuite par l'auteur et par d'autres. Pour un physicien, l'énergie libre F_N est définie par $F_N = (-1/\beta)p_N$. Le facteur $-1/\beta$ n'est pas utile au mathématicien, mais il ne semble pas judicieux d'augmenter à plaisir les difficultés de communication entre mathématiciens et physiciens. Le facteur $1/\beta$ est d'ailleurs nécessaire pour que F_N ait la dimension d'une énergie. On s'efforcera de calculer la limite de F_N quand $N \to \infty$. Si cette limite vaut, disons, a, cela signifie simplement

que pour N grand on a $Z_N \sim \exp(-\beta N a)$. L'énergie libre possède d'ailleurs un autre intérêt, car si on dérive $p_N = -\beta F_N$ par rapport à un paramètre au système (pour l'instant β est le seul paramètre, mais on en ajoutera bientôt d'autres), on voit apparaître Z_N au dénominateur, et des moyennes par rapport à la mesure de Gibbs, qui sont des quantités physiquement mesurables ; par exemple

$$\frac{\partial}{\partial\beta}p_N(\beta) = E\frac{1}{NZ_N}\sum_{\boldsymbol{\sigma}} -H_N(\boldsymbol{\sigma})\exp(-\beta H_N(\boldsymbol{\sigma})) = \frac{1}{N}E\langle -H_N(\boldsymbol{\sigma})\rangle \tag{6.2}$$

est l'espérance de l'opposé de l'énergie moyenne d'une configuration, où l'on a noté par $\langle\cdot\rangle$ la moyenne pour la mesure de Gibbs.

La première remarque à propos de (6.1) est l'utilisation de l'inégalité de Jensen

$$p_N = E\frac{1}{N}\log Z_N \leq \frac{1}{N}\log EZ_N = \log 2 + \frac{\beta^2}{4}\frac{(N-1)}{N} \tag{6.3}$$

par un calcul immédiat. Il devrait être évident que cette borne ne peut être du bon ordre pour β grand, le membre de gauche pouvant croître au plus comme β et non comme β^2.

THÉORÈME 6.1 [A-L-R].— *Pour $\beta \leq 1$, on a $\lim_{N\to\infty} p_N(\beta) = \log 2 + \beta^2/4$.*

L'article de Aizenmann, Lebowitz, Ruelle, l'un des premiers travaux rigoureux sur le modèle SK, reste très instructif. (Voir aussi [F-Z2] et surtout les profonds résultats de [F-Z1].) Il démontre beaucoup plus que le théorème 6.1, et en particulier un théorème limite central pour p_N (qui présente des fluctuations d'ordre 1). La démonstration repose sur des techniques de physique mathématique un peu lourdes. Une approche beaucoup plus élégante est proposée par Comets et Neveu qui utilisent cette fois les outils du calcul stochastique pour démontrer de nouveaux théorèmes limite centraux, toujours dans le domaine $\beta < 1$. Ce travail remarquable identifie des outils efficaces, mais dont l'utilité ne semble hélas pas s'étendre au-delà de quelques situations très particulières. Une troisième méthode, plus courte mais ne donnant que le théorème 6.1 (et pas les théorèmes limite centraux plus précis) est finalement introduite en [T2, section 12].

La condition $\beta \leq 1$ du théorème 6.1 est-elle optimale ? Cela est montré par la borne suivante.

THÉORÈME 6.2. [C] (Voir aussi [K-T-J]).— *Pour $\beta \geq 1$, on a*

$$\limsup_{N\to\infty} p_N \leq \log 2 + \beta - \frac{3}{4} - \frac{1}{2}\log\beta.$$

Que se passe-t-il pour $\beta > 1$?

C'est là le domaine des célèbres prédictions de Parisi. Les structures mathématiques remarquables inventées par Parisi sont aujourd'hui mieux comprises, en particulier grâce à des travaux récents de Bolthausen et Sznitman [B-S]. Quant à la question essentielle, c'est-à-dire de savoir si ces structures ont quoi que ce soit à voir avec le modèle SK, nous ne savons pratiquement rien de rigoureux à ce propos.

La situation est donc la suivante : si $\beta < 1$, le modèle est relativement bien compris (quoique bien des questions naturelles restent ouvertes). Si $\beta > 1$, il est au contraire extrêmement difficile. C'est une situation peu propice au progrès.

Revenons sur le cas $\beta < 1$. L'inégalité de Tchébitchef montre que la v.a. positive Z_N ne peut pas être souvent beaucoup plus grande que EZ_N. Il y a donc là une "barrière naturelle" à la taille de Z_N. Le fait que, asymptotiquement, on ait $EN^{-1} \log Z_N \simeq N^{-1} \log EZ_N$ implique que pour chaque $a > 0$, lorsque $N \to \infty$, on a $Z_N \geq 2^{-aN} EZ_N$ avec une probabilité qui tend vers 1. Cela signifie donc, en un sens, que Z_N "est le plus grand possible". Il devrait sembler naturel qu'une telle situation soit bien plus simple que la situation où Z_N est typiquement bien inférieur à EZ_N, car quelle valeur Z_N va-t-il alors "choisir" ? Une fâcheuse coïncidence a voulu que le modèle SK, de loin le plus connu, se trouve à $\beta < 1$ dans cette situation exceptionnelle. Cela n'est pas du tout représentatif des autres situations que nous étudierons, et donne une idée totalement fausse des problèmes "à haute température".

Pour sortir de ce cas très particulier, nous allons introduire un terme supplémentaire dans l'hamiltonien, remplaçant (3.2) par

$$H_N(\boldsymbol{\sigma}) = -\frac{1}{\sqrt{N}} \sum_{1 \leq i < j \leq N} g_{ij}\sigma_i\sigma_j - h \sum_{i \leq N} \sigma_i. \tag{6.4}$$

Le dernier terme, extrêmement naturel du point de vue physique, représente un "champ magnétique externe", et introduit une préférence des spins +1 par rapport aux spins -1. Dès que $h \neq 0$ (et quel que soit $\beta > 0$) la majoration (6.3) par l'inégalité de Jensen n'est plus efficace. De plus, il semble qu'aucune des méthodes menant au théorème 6.1 ne soit utile. Le cas $h \neq 0, \beta$ petit apparaît donc comme un cas de difficulté intermédiaire entre les situations $h = 0, \beta < 1$ et $h = 0, \beta > 1$.

7. LE MODÈLE SK AVEC CHAMP EXTERNE ; LES FORMULES RÉPLIQUE-SYMÉTRIQUE ; LE THÉORÈME DE PASTUR ET SHCHERBINA

Le modèle SK avec hamiltonien (6.4) dépend maintenant de deux paramètres : β et h. Expliquons brièvement les prédictions de [M-P-V]. Considérons une v.a. g

normale standard, et $q\ (=q(\beta,h))\geq 0$ satisfaisant

$$q = E\mathrm{th}^2\beta(g\sqrt{q}+h) \tag{7.1}$$

(le problème de l'existence ou de l'unicité d'un tel nombre ne semble pas abordé dans [M-P-V]) et considérons

$$\theta = \theta(\beta,h) = \beta^2 E\frac{1}{\mathrm{ch}^4\beta(g\sqrt{q}+h)}. \tag{7.2}$$

La région où $\theta<1$ sera appelée la région "haute température" et la région $\theta>1$ la région "basse température", qui est le domaine des prédictions de Parisi. Pour $\theta<1$, la prédiction est que

$$\lim_{N\to\infty} p_N(\beta,h) = \frac{\beta^2}{4}(1-q)^2 + E\log(2\mathrm{ch}\beta(g\sqrt{q}+h)). \tag{7.3}$$

Il faut insister sur le caractère remarquable de cette formule, qui signifie que pour N grand la valeur typique de $N^{-1}\log Z_N$ est donnée par le membre de droite de (7.3), alors que $N^{-1}\log EZ_N$ vaut approximativement $\beta^2/4+\log(2\mathrm{ch}\,h)$. Les formules (7.1), (7.3) sont appelées les formules "réplique-symétrique" du nom de la méthode (à laquelle nous avons déjà fait allusion) qui a permis de les découvrir, et qui est utilisée maintenant de façon systématique en physique. Devant le caractère mystérieux de cette méthode, Pastur et Shcherbina ont voulu donner une première justification de ces formules. Ils partent pour cela d'une des prédictions les plus remarquables de [M-V-P]. Considérons la v.a. $\overline{q}_N = N^{-1}\sum_{i\leq N}\langle\sigma_i\rangle^2$; alors [M-V-P] prédisent qu'à haute température la variable $\overline{q}_N$ est essentiellement constante ($\mathrm{Var}\,\overline{q}_N = E(\overline{q}_N - E\overline{q}_N)^2 \to 0$) mais que cela est *faux* à basse température, contrairement à l'intuition qu'une certaine "loi des grands nombres" devrait jouer. Remettant à plus tard la question de savoir pour quelles valeurs de β,h, on a bien $\mathrm{Var}\,\overline{q}_N\to 0$, Pastur et Shcherbina ont l'idée de montrer les formules réplique-symétrique sous cette condition. Personne hélas ne sachant faire cela avec l'hamiltonien (6.4), ils utilisent à la place un hamiltonien assez spécial et un peu mystérieux [P-S]. Ce travail est repris plus tard par Shcherbina, qui utilise cette fois l'hamiltonien

$$-\frac{1}{\sqrt{N}}\sum_{i<j} g_{ij}\sigma_i\sigma_j - h\sum_{i\leq N}\sigma_i h_i, \tag{7.4}$$

où les h_i sont une nouvelle suite indépendante de v.a. gaussiennes standard. On peut sans doute considérer comme secondaire le fait que la signification physique du

dernier terme de (7.4) ne soit pas aussi claire que celle du dernier terme de (7.5) ; mais au niveau technique la présence de ces nouvelles variables joue un rôle essentiel. Une version de ces résultats existe maintenant [T3, théorème 1.11] pour l'hamiltonien (4.1), en remplaçant $\overline{q}_N$ (qui vaut alors zéro) par $q_N^* = N^{-2}\sum_{i<j}\langle\sigma_i\sigma_j\rangle^2$.

Ce type de résultat a semble-t-il surtout un intérêt historique, puisque l'on sait maintenant aborder le vrai problème et montrer dans plusieurs situations bien différentes la validité des formules réplique-symétrique. Les trois sections suivantes tentent d'illustrer quelques idées de ces résultats à l'aide d'un schéma de preuve de (7.3) dans le domaine $\beta < \beta_0$, h quelconque.

8. LA MÉTHODE DE LA CAVITÉ ; LA NOTION D'ÉTAT PUR

Pour étudier le système gouverné par (6.3), il est naturel de tenter l'induction sur N, méthode connue en physique sous le nom de la méthode de cavité. (Pour se ramener de $N+1$ à N, on enlève un individu, ce qui crée une cavité...). Posons $N' = N+1$, considérons une nouvelle suite $(g_i)_{i\le N}$ indépendante de variables gaussiennes standard, qui soit indépendante de la suite $(g_{ij})_{i<j\le N}$. Pour $\boldsymbol{\sigma} \in \Sigma_N, \sigma_{N+1} \in \{-1,1\}$, considérons la quantité

$$H_{N'}(\boldsymbol{\sigma},\sigma_{N+1}) = \sqrt{\frac{N}{N'}}H_N(\boldsymbol{\sigma}) - \frac{1}{\sqrt{N'}}\sum_{i\le N} g_i\sigma_i\sigma_{N+1} - h'\sigma_{N+1}, \tag{8.1}$$

où $h = \sqrt{N/N'}h$, et où bien sûr $H_N(\boldsymbol{\sigma})$ est donné par (6.4). Si l'on pose $g_i = g_{iN+1}$, et que l'on identifie $(\boldsymbol{\sigma},\sigma_{N+1})$ à un élément de Σ_{N+1}, on voit que (8.1) correspond exactement au membre de droite de (6.4), au remplacement près de N par $N' = N+1$ et de h par h'. Désignons par $\langle\cdot\rangle$ la moyenne pour la mesure de Gibbs sur Σ_N, à température inverse β, pour l'hamiltonien (6.4), et par $\langle\cdot\rangle'$ la moyenne pour la mesure de Gibbs sur Σ_{N+1}, à température inverse $\beta' = \beta\sqrt{N'/N}$, correspondant à l'hamiltonien (7.1). On a alors des identités du type

$$\langle\sigma_{N+1}\rangle' = \frac{Av\langle\sigma_{N+1}\mathcal{E}\rangle}{Av\langle\mathcal{E}\rangle} \tag{8.2}$$

où

$$\mathcal{E} = \exp\beta\sigma_{N+1}\left(\frac{1}{\sqrt{N}}\sum_{i\le N} g_i\sigma_i + h'\right),$$

et où Av désigne la moyenne obtenue entre les valeurs de $\sigma_{N+1} = 1$ et $\sigma_{N+1} = -1$. Le point fondamental dans la méthode de la cavité est que dans (8.2) les v.a. g_i sont

indépendantes du désordre intervenant dans $\langle\cdot\rangle$ (c'est-à-dire des variables $(g_{ij})_{i<j\leq N}$ ce qui permet de travailler "à (g_{ij}) fixées").

La formule (8.2) montre qu'il serait désirable de savoir calculer $\langle\mathcal{E}\rangle$, ou plus généralement $X = \langle \exp tN^{-1/2}\sum_{i\leq N} g_i\sigma_i\rangle$. Essayons d'être optimiste, et de nous guider sur le cas où la mesure de Gibbs serait une mesure produit. Dans ce cas on aurait

$$X \simeq Y_1 = C\exp\frac{t}{\sqrt{N}}\sum_{i\leq N} g_i\langle\sigma_i\rangle \tag{8.3}$$

où C est indépendante des g_i. La première réaction est de choisir C de sorte que $E_gX = E_gY_1$, où E_g dénote intégration en les $(g_i)_{i\leq N}$ seulement, mais il s'avère plus fructueux d'oser écrire

$$X \simeq Y = \exp\left(\frac{t}{\sqrt{N}}\sum_{i\leq N} g_i\langle\sigma_i\rangle + \frac{t^2}{2}(1-\overline{q}_N)\right) \tag{8.4}$$

où $\overline{q}_N = N^{-1}\sum_{i\leq N}\langle\sigma_i\rangle^2$. Pour justifier (8.4), on peut essayer de montrer que $E_g(X-Y)^2$ est petit. On développe donc le carré, et on calcule chacun des termes. Pour calculer E_gX^2, on utilise une idée très utile des physiciens, qui consiste à écrire

$$X^2 = \left\langle \exp\frac{t}{\sqrt{N}}\sum_{i\leq N} g_i\sigma_i\right\rangle^2 = \left\langle \exp\frac{t}{\sqrt{N}}\sum_{i\leq N} g_i(\sigma_i^1+\sigma_i^2)\right\rangle \tag{8.5}$$

La dernière intégrale est sur Σ_N^2, munie de $G_N^{\otimes 2}$, et la fonction intégrée est fonction d'un couple $(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2)$ de configurations. C'est tout simplement la formule $(EA)^2 = EAB$ où B est une copie indépendante de A. Les deux copies de (Σ_N, G_N) (avec le *même* désordre) utilisées dans le membre de droite de (8.5) s'appellent des répliques (= copies), ce qui ne doit pas prêter à confusion avec la "méthode des répliques" qui elle aussi utilise des répliques, mais dont le nombre $n\to 0$. On déduit de (8.5) que

$$E_gX^2 = \left\langle \exp t^2\left(1+\frac{1}{N}\sum_{i\leq N}\sigma_i^1\sigma_i^2\right)\right\rangle$$

et par un calcul similaire

$$E_gXY = \left\langle \exp t^2\left(1+\frac{1}{N}\sum_{i\leq N}\sigma_i^1\langle\sigma_i\rangle\right)\right\rangle$$

et bien sûr

$$E_g Y^2 = \exp t^2(1+\overline{q}_N).$$

On remarque donc que l'on a $E_g(X-Y)^2 \simeq 0$ dès que $\langle (\frac{1}{N}\sum_{i\leq N}\sigma_i^1\sigma_i^2 - \overline{q}_N)^2\rangle \simeq 0$ (ce qui implique que $\langle (\frac{1}{N}\sum_{i\leq N}\sigma_i\langle\sigma_i\rangle - \overline{q}_N)^2\rangle \simeq 0$ en intégrant en $\boldsymbol{\sigma}^2$ à l'intérieur du carré plutôt qu'à l'extérieur). Cette remarque amène à poser la définition suivante, qui naturellement a du sens dans un cadre bien plus général que celui du modèle SK.

DÉFINITION 8.1.— *On dira qu'un système (à une valeur donnée des paramètres) est dans un* état pur *si l'on a*

$$\lim_{N\to\infty} E\Big\langle \Big(\frac{1}{N}\sum_{i\leq N}\sigma_i^1\sigma_i^2 - \overline{q}_N\Big)^2\Big\rangle = 0. \tag{8.6}$$

On voit donc sous cette hypothèse, par le calcul précédent que (3.2) devient

$$\langle\sigma_{N+1}\rangle' \simeq \operatorname{th}\beta\Big(\frac{1}{\sqrt{N}}\sum_{i\leq N} g_i\langle\sigma_i\rangle + h\Big), \tag{8.7}$$

ce qui est bien sûr beaucoup plus satisfaisant.

Il semble naturel de *définir* la phase "haute température" d'un système par la condition (8.6).

Une version de l'idée principale (8.4) de cette section se trouve déjà dans un article très confidentiel de Shcherbina [Sh1]. La différence est toutefois que dans la situation de [Sh1] on sait déjà que $\operatorname{Var}\overline{q}_N \to 0$; alors que l'une des forces de (8.4) est précisément de ne pas supposer cela a priori. Shcherbina déduit (8.4) de la condition $\operatorname{Var}\overline{q}_N \to 0$ par des arguments dont le principe mérite d'être mentionné ici. La fonction $N^{-1}\log Z_N$ est une fonction convexe de h d'ordre 1, et ses fluctuations sont d'ordre $N^{-1/2}$ (concentration de la mesure). Les dérivées en h de cette fonction ont donc (tout au moins pour beaucoup de valeurs de h) des variances d'ordre $o(1)$. Le calcul de ces variances, au moyen d'intégrations par parties en les variables h_i, fournit alors, lorsque $\operatorname{Var}\overline{q}_N \to 0$, les relations nécessaires à la preuve de (8.4).

9. CALCUL DANS UN ÉTAT PUR

Nous repoussons à la section 10 l'examen de la question centrale, c'est-à-dire de savoir quand le système est dans un état pur, et, afin d'illustrer l'utilité de ce concept, poursuivons l'étude du modèle SK (en supposant qu'il se trouve dans un état pur,

et que β est assez petit). On voit à l'examen de (8.7), qu'il serait bien pratique si $\overline{q}_N = N^{-1}\sum_{i\leq N}\langle\sigma_i\rangle^2$ était essentiellement indépendante du désordre, c'est-à-dire si $\Delta_N = \mathrm{Var}\,\overline{q}_N \to 0$. Cela ne résulte pas de (8.6). L'idée naturelle est d'essayer de relier $\Delta_{N+1}(\beta', h')$ et $\Delta_N(\beta, h)$. Nous ne détaillerons pas ce point, mais une fois que l'on a (8.6) et les relations type (8.7), on peut montrer que (si β est assez petit) on a $\Delta_{N+1}(\beta', h') \leq A\Delta_N(\beta, h) + o(1)$ ou $A < 1$ et $o(1) \to 0$ d'où il s'ensuit par itération que pour un certain $\beta_0 > 0$ on a

$$\Delta_N(\beta, h) \to 0 \text{ si } \beta \leq \beta_o. \tag{9.1}$$

On peut alors renforcer (8.6) en

$$\lim_{N\to\infty} E\Big\langle \Big(\frac{1}{N}\sum_{i\leq N}\sigma_i^1\sigma_i^2 - q_N\Big)^2\Big\rangle \to 0 \tag{9.2}$$

où $q_N = q_N(\beta, h) = E\overline{q}_N$. Cette dernière condition exprime "que les recouvrements de deux configurations sont constants", c'est-à-dire que si l'on choisit $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2$ indépendamment pour la mesure de Gibbs, le "recouvrement" $N^{-1}\sum_{i\leq N}\sigma_i^1\sigma_i^2$ de $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2$ est essentiellement indépendant de $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2$, et du désordre. C'est ce qu'expriment certains des diagrammes de [M-P-V], p. 47.

On déduit de (8.7), (9.1) que

$$E(\langle\sigma_{N+1}\rangle'^2) = E\mathrm{th}^2\beta(g\sqrt{q}_N + h) + o(1) \tag{9.3}$$

où g est normale standard. Par symétrie, on a $E(\langle\sigma_{N+1}\rangle')^2 = q_{N+1}(\beta' h')$ et donc (9.3) donne

$$q_{N+1}(\beta', h') \simeq E\mathrm{th}^2\beta(g\sqrt{q_N(\beta, h)} + h)$$

d'où il n'est pas bien difficile de déduire que, pour $\beta \leq \beta_0$, on a $q_N(\beta, h) \to q(\beta, h)$ où $q = q(\beta, h)$ vérifie (7.1).

Comment alors calculer l'énergie libre ? La quantité h n'intervient que par le produit $\tilde{h} = \beta h$. Il est pratique de considérer $\beta, \tilde{h}$ comme les nouveaux paramètres. On a alors

$$\begin{aligned}\frac{\partial}{\partial\beta}(p_N(\beta, \tilde{h})) &= \frac{1}{N}\, E\langle -H_N(\sigma)\rangle \\ &= \frac{1}{N} E\frac{1}{\sqrt{N}}\sum_{i<j} g_{ij}\langle\sigma_i\sigma_j\rangle.\end{aligned}$$

On fait appel à la formule d'intégration par parties $E(g\varphi(g)) = E\varphi'(g)$ si g est $N(0,1)$, formule qui se révèle extrêmement utile, et qui donne ici après un petit calcul

$$\frac{\partial}{\partial\beta}(p_N(\beta,\tilde{h})) = \frac{\beta}{N^2}E\sum_{i<j}(1-\langle\sigma_i\sigma_j\rangle^2)$$
$$= \frac{\beta(N-1)}{2N} - \frac{\beta}{N^2}E\sum_{i<j}\langle\sigma_i\sigma_j\rangle^2.$$

Or, en faisant intervenir les répliques, $\langle\sigma_i\sigma_j\rangle^2 = \langle\sigma_i^1\sigma_j^1\sigma_i^2\sigma_j^2\rangle$, et donc

$$E\frac{1}{N^2}\sum_{i<j}\langle\sigma_i\sigma_j\rangle^2 = E\frac{1}{2}\Big(\Big\langle\Big(\frac{1}{N}\sum_{i\le N}\sigma_i^1\sigma_i^2\Big)^2\Big\rangle - \frac{1}{N}\Big)$$

et utilisant (9.2) et $q_N \to q$ cela donne

$$\lim_{N\to\infty}\frac{\partial}{\partial\beta}(p_N(\beta,\tilde{h})) = \frac{\beta}{2}(1-q^2)$$

et, puisque (7.3) est correcte à $\beta = 0$, il suffit alors pour la vérifier de montrer que la dérivée de son membre de droite par rapport à β est $\beta(1-q^2)/2$.

Quelle sera la structure de la mesure de Gibbs ? Le point un peu subtil est que globalement (même dans le cas $h = 0$) G_N est très différente d'une mesure produit (et ceci quoique certaines de ces propriétés globales rappellent les mesures produit, comme le fait que le recouvrement de deux configurations soit essentiellement indépendant de ces configurations), mais que "localement" (c'est-à-dire si l'on ne regarde qu'un nombre fini de spins) elle ressemble à une mesure produit. Pour étudier ce phénomène, retournons vers la relation (8.6). Celle-ci est mal commode à cause du terme $\overline{q}_N$. On la transforme à l'aide de l'idée connue en probabilités sous le nom de symétrisation, pour en déduire (par différence), en considérant maintenant trois répliques $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \boldsymbol{\sigma}^3$, que

$$\lim_{N\to\infty} E\Big\langle\Big(\frac{1}{N}\sum_{i\le N}(\sigma_i^1-\sigma_i^2)\sigma_i^3\Big)^2\Big\rangle = 0 \tag{9.4}$$

puis que

$$\lim_{N\to\infty} E\Big\langle\Big(\frac{1}{N}\sum_{i\le N}(\sigma_i^1-\sigma_i^2)(\sigma_i^3-\sigma_i^4)\Big)^2\Big\rangle = 0. \tag{9.5}$$

(Il n'est pas difficile en fait de montrer que (8.6), (9.4), (9.5) sont équivalentes.) Si l'on écrit $c_i = (\sigma_i^1 - \sigma_i^2)(\sigma_i^3 - \sigma_i^4)$ et que l'on développe le produit $(\sum_{i\leq N} c_i)^2$, on voit par symétrie sur les coordonnées que (9.5) équivaut à $E\langle c_1 c_2\rangle \to 0$, ou encore à $E(\langle\sigma_1\sigma_2\rangle - \langle\sigma_1\rangle\langle\sigma_2\rangle)^2 \to 0$, c'est-à-dire au fait que deux spins différents ne sont pas asymtotiquement corrélés pour la mesure de Gibbs. (Le lecteur ne manquera pas d'observer que cette même méthode appliquée à (9.4) est moins intéressante !) En remarquant que $|\sum_{i\leq N} c_i| \leq 2N$, et donc que d'après (9.4) on a $\lim_{N\to\infty} E\langle(N^{-1}\sum_{i\leq N} c_i)^n\rangle \to 0$ pour chaque n, on obtient avec un peu plus de travail la proposition suivante, valable en toute généralité et pas seulement pour le modèle SK.

PROPOSITION 9.1.— *La condition* (8.4) *est équivalente au fait suivant : Pour chaque p, quand $N \to \infty$, la projection $G_{N,p}$ de G_N sur $\{-1,1\}^p$ (c'est-à-dire la loi sous G_N de $\boldsymbol{\sigma} \to (\sigma_1,...,\sigma_p)$) est asymptotiquement une mesure produit, au sens que l'espérance de sa distance (en quelque sens raisonnable que l'on choisisse) à une mesure produit tend vers zéro.*

Bien sûr $G_{N,p}$ est aléatoire. Comment dépend-elle du désordre ? Une mesure sur $\{-1,1\}$ est déterminée par son barycentre, et donc $G_{N,p}$ est (asymptotiquement) déterminée par $(\langle\sigma_i\rangle)_{i\leq p}$. Quelle est la structure de cette suite de v.a. ? Une fois encore, la réponse est en un sens la plus simple possible. En utilisant non pas seulement (8.4) mais aussi (9.2) et le fait que $q_N^{-1} \to q$, et en exprimant les spins $\sigma_1,...,\sigma_p$ en fonction des autres (par une généralisation de la méthode de la cavité), on obtient par un schéma de preuve qui s'avère très général le résultat suivant, qui donne une bonne compréhension "d'un nombre fini de spins".

PROPOSITION 9.2.— *Pour le modèle SK, si $\beta < \beta_0$, h quelconque, si on fixe p, les variables $(\langle\sigma_i\rangle)_{i\leq p}$ sont, lorsque $N \to \infty$, asymptotiquement indépendantes de même loi, et cette loi est celle de* $\operatorname{th}\beta(g\sqrt{q}+h)$ *où g est $N(0,1)$ et où q est donné par* (7.1).

10. MONTRER QUE LE SYSTÈME EST DANS UN ÉTAT PUR ET LE PROBLÈME CENTRAL DE LA HAUTE TEMPÉRATURE

Il existe un argument assez simple [T1, théorème 3.1] montrant que le modèle SK est dans un état pur à haute témperature. Cet argument semble difficile à étendre à d'autres modèles, et nous préférons esquisser une approche plus générale. Intéressons-nous à la quantité $D_N = D_N(\beta,h)$ dont (9.4) considère la limite. On va tenter de relier $D_{N+1} = D_{N+1}(\beta',h')$ à $D_N(\beta,h)$, où l'on utilise les notations de la section 8.

Une bonne façon de commencer le calcul est d'écrire

$$D_{N+1} = E\Big\langle \Big(\frac{1}{N+1}\sum_{i\le N+1}(\sigma_i^1-\sigma_i^2)\sigma_i^3\Big)^2\Big\rangle'$$
$$= E\Big\langle (\sigma_{N+1}^1-\sigma_{N+1}^2)\sigma_{N+1}^3\sum_{i\le N+1}(\sigma_i^1-\sigma_i^2)\sigma_i^3\Big\rangle'$$

en utilisant la symétrie entre les coordonnées, d'où

$$D_{N+1} = \frac{1}{N+1}E\big\langle((\sigma_{N+1}^1-\sigma_{N+1}^2)\sigma_{N+1}^3)^2\big\rangle' + \frac{N}{N+1}E\big\langle(\sigma_{N+1}^1-\sigma_{N+1}^2)\sigma_{N+1}^3 f\big\rangle' \tag{10.1}$$

et

$$D_{N+1} \le \frac{4}{N+1} + \frac{N}{N+1}E\big\langle(\sigma_{N+1}^1-\sigma_{N+1}^2)\sigma_{N+1}^3 f\big\rangle' \tag{10.2}$$

où $f = N^{-1}\sum_{i\le N}(\sigma_i^1-\sigma_i^2)\sigma_i^3$. On relie alors $\langle\cdot\rangle'$ à $\langle\cdot\rangle$ dans le dernier terme, comme dans (8.2).

La difficulté consiste à trouver une méthode pour estimer ce terme, qui fait intervenir G_N, sur laquelle on sait très peu de choses. Il n'est malheureusement pas possible de donner des détails sur ce point très technique. On effectue une sorte de développement limité, qui peut se faire avec différents niveaux de précision, mais dans lequel il convient bien sûr de contrôler les termes d'erreur. On peut contrôler ceux-ci, pour $\beta \le 1$, par $L\beta D_{N,2}$, où L est un nombre, et où

$$D_{N,n} = E\Big\langle\Big(\frac{1}{N}\sum_{i\le N}(\sigma_i^1-\sigma_i^2)\sigma_i^3\Big)^{2n}\Big\rangle. \tag{10.3}$$

On a bien sûr $D_{N,n} \le 2^{n-1}D_N$. Si β est petit, le premier objectif étant de montrer à l'aide de (10.2) que

$$D_{N+1} \le \frac{4}{N+1} + BD_N \tag{10.4}$$

ou $B < 1$ (ce qui permet de conclure que $D_N \to 0$ par itération, et même de montrer que ND_N reste borné), ces termes d'erreurs ne seront pas trop génants. Mais, si l'on veut aller plus loin, par exemple calculer $\lim_{N\to\infty} ND_N$, il faut alors montrer

que $D_{N,2} \ll D_N$. Or, il n'est pas évident que c'est le cas, même si $D_N \ll 1$. Il est même possible de montrer que pour certains modèles (et en particulier le modèle à "p spins" que nous considérerons plus tard) il se peut très bien que $D_N \ll 1$ mais que $D_{N,2}$ et D_N soient du même ordre. Cela oblige à contrôler $D_{N,2}$ par induction sur N, et de proche en proche chaque $D_{N,n}$. En suivant cette philosophie, on peut alors (pour β petit) obtenir une description très précise (par exemple calculer chaque $\lim_{N\to\infty} N^r D_{N,r}$) et en fait la précision que l'on peut obtenir ne semble plus limitée que par l'énergie que l'on veut bien y consacrer.

Une approche assez différente des formules (7.1), (7.3) a été proposée par M. Shcherbina [Sh2]. Cette approche ne s'applique qu'à l'hamiltonien (7.4), car elle repose sur les techniques spéciales permises par cet hamiltonien qui ont été évoquées à la fin de la section 8. Shcherbina ne dégage pas la notion d'état pur, mais à la place de D_N utilise la quantité $\Delta_N = \operatorname{Var} \overline{q}_N$. Son but est de montrer que $\lim \Delta_N = 0$, car les formules réplique-symétrique résulteront alors de [Sh1]. L'aspect remarquable du travail de Shcherbina est que celle-ci dégage sans information a priori sur la mesure de Gibbs, une relation compliquée mais *exacte* sur Δ_N dont on peut extraire dans certaines conditions le fait que $\Delta_N \to 0$. Cette information est obtenue par un véritable tour de force, à l'aide de certaines intégrations par parties choisies judicieusement parmi beaucoup de choix possibles. Il serait évidemment très intéressant de comprendre s'il y a là une sorte de coïncidence fortuite ou plutôt l'amorce d'une méthode générale. Il faut également signaler que la méthode de [Sh2] semble incapable d'obtenir des résultats très précis tels par exemple que le calcul de $\lim_{N\to 0} N D_N$.

Comment les physiciens justifient-ils la prédiction que la zone à haute température correspond à $\theta < 1$ (ou θ est donné par (7.2)) ? Si $C_N = E(\langle \sigma_1, \sigma_2 \rangle - \langle \sigma_1 \rangle \langle \sigma_2 \rangle)^2$, ils arrivent, semble-t-il en supposant

$$E\Big\langle \Big(\frac{1}{N} \sum_{i \le N} \sigma_i^1 \sigma_i^2 - q \Big)^2 \Big\rangle \to 0 \tag{10.5}$$

où q vérifie (7.1), et sans contrôle de certains termes d'erreurs à la relation

$$C_N = \theta C_N + A \tag{10.6}$$

où $A = E(4/\mathrm{ch}^2 \beta(g\sqrt{q}+h))$ est positif, et en concluent donc que la condition cherchée est $\theta < 1$. Cet argument paraît plus "solide" dans le cas $\theta > 1$ que dans le cas $\theta < 1$, où il est au plus une vérification du type "consistance interne". Mais dans tous les cas, le problème des termes d'erreur indûment négligés dans (10.6) est crucial, puisque

nous avons vu (dans le cas de $D_{N,2}$ et D_N) que des termes qui semblent "d'ordre supérieur" ne le sont pas nécessairement.

L'argument précédent a néanmoins pu être mis sous forme rigoureuse si $\theta > 1$ dans [T5, Th.3.3] où il est montré que dans ce cas le système manifeste des instabilités frappantes ("des causes microscopiques ayant des effets macroscopiques") qui peuvent s'interpréter comme une violation d'une version forte de (8.5).

Le problème ouvert central semble maintenant de démontrer la validité des formules réplique-symétrique dans toute la zone $\theta < 1$. Cela n'a pas une importance considérable en soi en ce qui concerne le modèle SK. Mais celui-ci est (à haute température) de loin le plus simple des modèles que nous considérerons et il n'y a pas d'espoir de contrôler toute la zone à haute température pour les autres modèles si nous ne savons pas le faire pour SK. Or, pour certains de ces modèles, il y a un enjeu réel à contrôler toute cette zone, qui dans quelques cas pourrait (comme le conjecturent les physiciens) s'étendre jusqu'à la température zéro et permettre ainsi une information précise sur les problèmes d'optimisation combinatoire dont ces modèles sont issus par la démarche de la section 4. La difficulté principale peut se résumer approximativement ainsi. Il semble impossible de contrôler D_N sans contrôler tous les $D_{N,n}$. Tout calcul de $D_{N,n}$ semble nécessiter des termes d'erreur du type $BD_{N,n+1}$ (ou B est indépendant de N et borné inférieurement si β ne s'approche pas de zéro) alors qu'il ne faudrait pas des termes d'erreur plus grands que $(1-\theta)D_{N,n}$ (ce qui peut se comprendre au vu de la situation semblable de (10.6)). Mais si n est d'ordre $N, D_{N,n+1}$ et $D_{N,n}$ doivent être du même ordre (par analogie avec les moments des variables gaussiennes) et donc $BD_{N,n+1}$ est plus grand que $(1-\theta)D_{N,n}$.

11. INTERSECTIONS D'ENSEMBLES ALÉATOIRES

11.1. Généralités

Nous avons acquis sur le modèle SK une philosophie d'attaque : Montrer successivement (8.6) (ou plutôt (9.4)), puis (9.2), et enfin déterminer la limite de q_N. (Un physicien par contre considérera que (10.4) va de soi et n'aura donc pour toute tâche que de déterminer q.) Il est donc temps d'étudier d'autres modèles. Dans cette section, nous considérerons deux problèmes de même structure. Nous aurons une famille $(A_k)_{k\leq M}$ aléatoire indépendante équidistribuée de sous-ensembles de Σ_N, et il faudra étudier $\lambda_N(\cap_{k\leq M} A_k)$, où λ_N dénote la mesure uniforme de Σ_N. La difficulté sera que M est de l'ordre N, et donc que si $a = E\lambda_N(A_k) < 1$, on a $E\lambda_N(\cap_{k\leq M} A_k) = a^M$, ce qui est extrêmement petit et rend donc difficile de déterminer si $\cap_{k\leq M} A_k$ est ou non

vide. Le problème est bien sûr que typiquement $\lambda_N(\cap_{k\leq M} A_k)$ est bien plus petit que son espérance. Une difficulté supplémentaire est que $\lambda_N(\cap_{k\leq M} A_k)$ peut en principe dépendre essentiellement de chacun des ensembles A_k. Il est donc naturel d'essayer d'abord un problème plus facile, et de considérer l'hamiltonien

$$H_{N,M}(\boldsymbol{\sigma}) = -\sum_{k\leq M} 1_{A_M}(\boldsymbol{\sigma}) \tag{11.1}$$

qui compte le nombre d'ensembles A_k auxquels appartient $\boldsymbol{\sigma}$. L'hamiltonien (11.1) "favorise" les configurations qui appartiennent à beaucoup d'ensembles A_k. Après introduction d'une température, il permet l'utilisation des idées de la mécanique statistique, et l'on peut, comme en section 4, espérer que le calcul de l'énergie libre à β grand permettra (après un travail supplémentaire) de résoudre le problème initial. On se fixe un nouveau paramètre α, et l'on considère dans la suite le cas où $M = \lfloor \alpha N \rfloor$.

11.2. La capacité du perceptron binaire

Dans ce problème [G2], d'intérêt considérable en théorie des réseaux de neurones, les ensembles A_k sont des demi-espaces à direction aléatoire, à distance fixe de l'origine, c'est-à-dire

$$A_k = \{\boldsymbol{\sigma} \in \Sigma_N; \sum_{i\leq N} \xi_i^k \sigma_i \geq b\sqrt{N}\}$$

où b est un nombre, et où les variables ξ_i^k sont indépendantes et $P(\xi_i^k = \pm 1) = 1/2$. (Ce qui est le modèle naturel de direction aléatoire pour les applications aux réseaux de neurones.) Étant donnés α, β, on va chercher à calculer l'énergie libre correspondant à l'hamiltonien (11.1), suivant le schéma évoqué au début de ce paragraphe. Apparaissent dès l'abord des difficultés qui n'étaient pas présentes dans le cas du modèle SK. Tout d'abord, afin de relier le système à $N+1$ spins au système à N spins, et afin de faire des développements limités, il semble indispensable de remplacer dans (11.1) la quantité $1_{A_k}(\boldsymbol{\sigma})$ par $u(N^{-1}\sum_{i\leq N} \xi_i^k \sigma_i)$ où u est une fonction suffisamment différentiable. Les termes principaux de ces développements limités (c'est-à-dire ceux qui ne tendent pas vers zéro avec N) dépendent de u', u'', mais afin de pouvoir, par un argument limite, retrouver la valeur de l'énergie libre correspondant à l'hamiltonien (11.1), il est indispensable de contrôler les termes d'erreur en fonction seulement de $\sup|u|$ et non de $\sup|u'|, \sup|u''|$. Une autre difficulté qui apparaît, semble-t-il, dans tous les modèles autres que le modèle SK est que, lorsque l'on essaye de calculer D_{N+1} en fonction de D_N, on voit apparaître de nouvelles quantités qui ne semblent pas reliées simplement à D_N, et qui doivent à leur tour être contrôlées. Il est très

utile, suivant une idée de Mézard [M], d'utiliser non seulement l'induction sur N mais également l'induction sur M. L'énergie libre est maintenant rigoureusement connue si $L\alpha \exp L\beta \leq 1$, où L est un nombre [T6].

11.3. Le problème du p-sat

On veut étudier ici un modèle aléatoire d'un célèbre problème d'informatique théorique. On se fixe un entier p, et les ensembles (A_k) sont alors des sous-cubes aléatoires de Σ_N, c'est-à-dire du type

$$\{\boldsymbol{\sigma} \in \Sigma_N; \forall \ \ i \in I, \sigma_i = \eta_i\}, \tag{11.2}$$

où $\operatorname{card} I = p, \eta_i = \pm 1$, tous les choix possibles de I et de η_i' étant équiprobables. Il y a une différence fondamentale entre ce modèle et les autres modèles considérés ici. Il n'y a typiquement qu'un nombre borné indépendamment de N (en probabilité) d'ensembles I_k contenant un individu donné i, où bien sûr I_k correspond à A_k par (11.2). Cela signifie qu'un spin "n'interagit qu'avec un nombre fini d'autres spins" et qu'il ne faut donc plus s'attendre à des effets type théorème limite central qui expliquent l'apparition des variables gaussiennes dans des formules telles que (7.3). Le modèle est alors intrinsèquement compliqué. Les physiciens décrivent cela en disant que le "paramètre du modèle SK est un nombre" (c'est-à-dire q, qui est calculé par (7.1)) mais "que le paramètre du modèle du p-sat est une fonction" (la loi limite de $\langle \sigma_1 \rangle$, qui obéit alors à une équation fonctionnelle compliquée). Il est intéressant que malgré cela le schéma d'approche de 11.1 soit encore approprié. Il faut toutefois inventer une nouvelle méthode pour conduire les calculs. Pour expliquer de façon imagée la difficulté, disons que les termes d'erreur ont une forte tendance, si l'on n'est pas très prudent, à être du type $\sqrt{D_N}$, ce qui conduit à des relations de forme $D_{N+1} \leq A\sqrt{D_N} + o(1)$, lesquelles ne permettent pas de conclure que $D_N \to 0$. Quelle que soit la valeur de α, l'énergie libre est maintenant rigoureusement connue si β est assez petit [T6].

12. LE MODÈLE DE HOPFIELD ; LES TRAVAUX DE BOVIER, GAYRARD, PICCO

Ce modèle est devenu célèbre en tant que modèle de mémoire. Le désordre y est représenté par une suite ξ_i^k définie comme en 11.2, et l'hamiltonien est donné par

$$H_{N,M}(\boldsymbol{\sigma}) = -\frac{1}{2N} \sum_{k \leq M} \Big(\sum_{i \leq N} \xi_i^k \sigma_i \Big)^2. \tag{12.1}$$

Une façon de comprendre cette formule est d'observer que le second membre est un bon candidat à être une fonction de $\boldsymbol{\sigma}$ prenant une valeur fortement négative dès que $\boldsymbol{\sigma} = \boldsymbol{\xi}^k = (\xi_i^k)_{i \leq N}$, puisque la contribution du terme correspondant est alors $-N/2$. L'idée est que le système "garde en mémoire" chaque configuration $\boldsymbol{\xi}^k$. Celles-ci seront appelées prototypes. Le cas le plus intéressant, le seul qui sera considéré ici, est celui où $M = \lfloor \alpha N \rfloor$, une proportion fixe de N. L'étude du "paysage" d'énergie donnée par (12.1) est alors difficile, pour les raisons exposées avant l'énoncé du problème 3.1. Il est toutefois connu que si $\alpha \leq 0,05$, il existe pour N grand un minimum local de $H_{N,M}$ au voisinage (au sens de la distance de Hamming) de chaque prototype [N], et de beaux travaux de D. Loukianova [L] précisent ce résultat. Un problème ouvert particulièrement simple est la conjecture suivante. Si $\alpha = 0,1$, pour N grand, il existe au voisinage de chaque prototype un minimum local, mais le minimum global de H_N est plus profond. La difficulté tient bien sûr au fait que nous savons très mal estimer ce dernier.

La fonction $H_{N,M}(\boldsymbol{\sigma})$ est relativement bien comprise pour α petit. Non seulement elle présente un "puits" au voisinage de chaque prototype, mais son minimum global se trouve dans l'un de ces puits. Par contraste, la situation est beaucoup plus compliquée pour le modèle SK où les simulations numériques indiquent la présence d'un grand nombre de "puits" étroits et profonds, dont la position est très mal comprise. Il est donc naturel de penser a priori que le modèle de Hopfield doit être particulièrement simple à étudier. Il se trouve que ce n'est pas le cas, mais c'est sans doute cela (et son rôle de modèle de mémoire) qui a poussé Bovier et Gayrard (parfois en collaboration avec Picco) à lui consacrer de nombreux travaux. (Voir [B-G] et les références incluses.) Nous ne discuterons que le cas $\beta > 1$ (*basse* température), le plus intéressant. Dans ce cas, la mesure de Gibbs a tendance à se concentrer dans les puits de basse énergie autour des prototypes. Afin d'éviter qu'elle "n'hésite" entre ces puits, il est avantageux de favoriser l'un deux, d'introduire un nouveau paramètre $h > 0$ et de considérer l'hamiltonien

$$H_{N,M}(\boldsymbol{\sigma}) = -\frac{1}{2N} \sum_{k \leq M} \Big(\sum_{i \leq M} \xi_i^k \sigma_i \Big)^2 - h \sum_{i \leq N} \xi_i^1 \sigma_i. \tag{12.2}$$

Les auteurs de [B-G], [B-G-P] observent que $H_{N,M}(\boldsymbol{\sigma})$ étant défini en fonction des quantités $m_k(\boldsymbol{\sigma}) = N^{-1} \sum_{i \leq N} \xi_i^k \sigma_i$, il est naturel de considérer l'image $\overline{G}_N$ dans $\mathbb{R}^M$ de la mesure de Gibbs par l'application $\boldsymbol{\sigma} \to (m_k(\boldsymbol{\sigma}))_{k \in M}$. Ils montrent alors que si $\alpha \leq L^{-1} \min(1, (\beta - 1)^2)$, $\overline{G}_N$ est (avec une probabilité écrasante) essentiellement concentrée sur une petite boule centrée au point $m^* \boldsymbol{e}_1$, où $e_1 = (1, 0, ...)$ et où m^* est

donné par $\text{th}\,\beta(m^* + h) = m^*$. Leurs résultats sont qualitativement exacts, tant en ce qui concerne les valeurs de α que le rayon de la petite boule, et nécessitent un travail considérable. Il est possible que cette approche, cherchant à obtenir des résultats qualitatifs plutôt que des formules exactes, soit promise à beaucoup d'avenir, car il ne faut pas s'attendre à toujours disposer de formules exactes, et les raisons réelles de l'existence de celles-ci dans les modèles que nous considérons sont à vrai dire quelque peu mystérieuses.

La méthode de la cavité a maintenant permis une compréhension très précise du modèle de Hopfield dans le domaine de paramètre $\beta < 1, \alpha \leq L^{-1} \min(1/\log\beta, (\beta - 1)^2)$. Les arguments reprennent une fois encore l'approche de 11.1, mais nécessitent à nouveau de multiples développements techniques, et sont rendus extrêmement longs par la nécessité de contrôler simultanément de nombreuses quantités. Le travail remarquable de Bovier et Gayrard [B-G], utilisant en particulier un argument géométrique de convexité, et les inégalités de Brascamp-Lieb, permet toutefois de façon plus simple le calcul de l'énergie libre dans le sous-domaine des paramètres $\alpha \leq L^{-1} \min(\beta, (\beta - 1)^2)$.

Il faut souligner que les estimations "a priori" de [B-G-P] sur la "localisation" de $\overline{G}_N$, qui ont été évoquées plus haut, jouent un rôle essentiel dans les arguments de cavité. La méthode de la cavité peut être comparée à une induction : les arguments a priori permettent de l'amorcer. Si l'on préfère une autre analogie (moins superficielle qu'il n'y paraît) on voudrait montrer qu'une certaine application T d'un espace métrique compliqué X dans lui-même possède un unique point fixe. Les estimations a priori montrent qu'il suffit d'étudier un petit sous-ensemble de X, fixé par T. La méthode de la cavité montre que T est une contraction sur ce sous-ensemble (alors qu'elle ne l'est pas partout) et produit le point fixe par itération.

13. LE REM DE DERRIDA

Nous allons maintenant tenter de donner une (toute petite) idée des structures inventées par les physiciens à basse température, et des difficultés formidables auxquelles se heurte l'étude rigoureuse de cette phase. Nous commençons par un modèle très simple, mais très instructif, le "Random Energy Model" (REM) introduit par Derrida. L'idée en est aussi simple qu'efficace. Puisqu'une difficulté essentielle du modèle SK est que les énergies de deux configurations différentes sont corrélées de façon compliquée, considérons donc le modèle où les énergies $H_N(\boldsymbol{\sigma})$ sont des gaussiennes indépendantes, avec $EH_N^2(\boldsymbol{\sigma}) = N$ (pour avoir la bonne normalisation). Cela n'a

bien sûr pas de prétention d'être un modèle réaliste, mais il est remarquablement intéressant. Il n'est pas très difficile de montrer que $\lim_{N\to\infty} p_N$ vaut $\log 2 + \beta^2/2$ si $\beta \leq \sqrt{2\log 2}$ et vaut $\beta\sqrt{2\log 2}$ si $\beta \geq \sqrt{2\log 2}$.

Le plus intéressant est toutefois la structure de la mesure de Gibbs lorsque $\beta > \sqrt{2\log 2}$. Expliquons ce qui se passe, sans donner de détails techniques. Puisqu'à $\boldsymbol{\sigma}$ donné, $H_N(\boldsymbol{\sigma})$ est normale centrée de variance N, l'espérance du nombre d'énergies $H_N(\boldsymbol{\sigma})$ qui se trouvent dans un petit intervalle de longueur dx autour de $x - N\sqrt{2\log 2}$ vaut asymptotiquement

$$\frac{2^N dx}{\sqrt{2\pi}} \exp -\frac{(x - N\sqrt{2\log 2})^2}{2N} \approx \frac{dx}{\sqrt{2\pi}} \exp x\sqrt{2\log 2}.$$

Il y a d'autre part indépendance aux voisinages de deux points différents. Puisque cela ne change rien à la mesure de Gibbs d'ajouter la constante $N\sqrt{2\log 2}$ à chaque énergie, on peut visualiser les plus basses de ces énergies $H_N(\boldsymbol{\sigma})$ comme provenant d'un processus de Poisson ponctuel dont la mesure d'intensité a une densité $(2\pi)^{-1/2} \exp x\sqrt{2\log 2}$ par rapport à la mesure de Lebesgue. Un calcul simple (changement de variable) montre alors que les plus grands des "poids" $\exp(-\beta H_N(\boldsymbol{\sigma}))$ peuvent asymtotiquement être vus comme réalisation d'un processus de Poisson ponctuel dont la mesure d'intensité a cette fois une densité $(1/\beta\sqrt{2\pi})x^{-m-1}$, où $m = \sqrt{2\log 2}/\beta$.

Cette représentation permet de visualiser très clairement la différence de comportement entre haute température ($\beta < \sqrt{2\log 2}, m > 1$) et basse température ($\beta > \sqrt{2\log 2}, m < 1$). Désignons par $(x_\alpha)_{\alpha\geq 1}$ (où $x_1 \geq x_2 \geq ...$) les points produits par une réalisation du processus précédent. A haute température l'intégrale $\int_{0+} x^{-m} dx$ est infinie, et la somme $S = \sum_{\alpha\geq 1} x_\alpha$ est infinie p.s. ce qui signifie que même les plus grands des poids $\exp(-\beta H_N(\boldsymbol{\sigma}))$ sont négligeables devant la somme Z_N de tous ces poids, et donc que la mesure de Gibbs ne charge asymptotiquement aucune configuration. Au contraire à basse température $m < 1$ et S est finie p.s. La mesure de Gibbs charge alors des configurations. Si l'on ordonne celles-ci par ordre décroissant de leurs poids pour cette mesure, la distribution de ces poids converge vers celle de la suite $(x_\alpha/S)_{\alpha\geq 1}$ lorsque $N \to \infty$. Cette dernière distribution (dépendant du seul paramètre m) est tout à fait naturelle en probabilité [P-Y]. Les physiciens prédisent que cette famille de distributions possède un caractère universel dans les problèmes à basse température.

14. LE MODÈLE DES INTERACTIONS À p-SPINS

Ce modèle (également inventé par Derrida) peut être présenté comme une généralisation du modèle SK. Les interactions de chaque paire sont remplacées par les interactions de chaque p-uples, et l'hamiltonien est donné par

$$H_N(\sigma) = -\left(\frac{p!}{2N^{p-1}}\right)^{1/2} \sum g_{i_1 \ldots i_p} \sigma_{i_1} \ldots \sigma_{i_p}. \tag{14.1}$$

La sommation est effectuée sur tous les choix de $i_1 < \ldots < i_p$, et les v.a. $g_{i_1 \ldots i_p}$ sont indépendantes gaussiennes standard. Pour $p = 2$, c'est le modèle SK, mais pour $p > 2$ le comportement est très différent. L'idée essentielle de (14.1) est que (par un calcul facile)

$$EH_N(\sigma)H_N(\sigma') \simeq \frac{N}{2}\Big(\frac{1}{N}\sum_{i \leq N} \sigma_i \sigma_i'\Big)^p$$

et donc que, plus p est grand, moins les énergies de configurations différentes sont corrélées, et donc plus le modèle "approche" le REM de Derrida.

Il existe une valeur critique β_p telle que $\lim_{N \to \infty} p_N(\beta) = \beta^2/4$ si $\beta \leq \beta_p$, mais $\limsup_{N \to \infty} p_N(\beta) < \beta^2/4$ si $\beta > \beta_p$. Pour $\beta > \beta_p$ (mais pas trop grand car apparaissent alors des phénomènes compliqués) les physiciens prédisent une situation simple et remarquable : le système se décompose en "états" (il n'est malheureusement semble-t-il jamais donné de définition mathématiquement précise de ce que cela signifie...) et la distribution des poids relatifs de ceux-ci pour la mesure de Gibbs est gouvernée (à la limite) par la distribution décrite en section 14, où le paramètre m dépend bien sûr de β et p. Le caractère remarquable de cette "décomposition en états" est que l'hamiltonien ne semble avantager aucune configuration, et l'on voit mal comment ces états émergent de celui-ci. (Par contraste, l'hamiltonien du modèle de Hopfield avantage clairement les prototypes). Nous proposerons plus tard une interprétation de ce phénomène. La décomposition en états a été partiellement justifiée rigoureusement, mais un énoncé précis nous entraînerait dans une discussion trop longue [T6]. Cette justification utilise une méthode indirecte, qui ne donne pas d'information sur les poids respectifs des "états" pour la mesure de Gibbs. Il semble qu'obtenir des informations rigoureuses sur ceux-ci soit maintenant la clef de progrès ultérieurs sur la phase à basse température. L'une des difficultés est que la valeur du paramètre m prévue par les physiciens n'est pas donnée par une condition simple obtenue par cavité (telle la condition (7.1) qui détermine q) mais au contraire est déduite du principe physique général suivant "si la structure d'un système dépend d'un paramètre, celui-ci prend la valeur qui maximise l'énergie libre". (Il est intéressant de remarquer que la

valeur de q fournie par (7.1) est précisément celle qui minimise le membre de droite de (7.2), comme on doit s'y attendre d'après le principe précédent). Il sera bien évidemment plus difficile de faire apparaître le paramètre m, qui est "déterminé par un principe faible" que s'il était déterminé par une condition plus forte. On peut aussi remarquer que même si l'on savait que, pour chaque N, la distribution des poids des états ressemble à la distribution prévue pour une valeur $m(N)$ du paramètre m, la méthode de la cavité, à cause des petites erreurs qu'elle crée et qui s'accumulent à chaque étape, est intrinsèquement impuissante à démontrer l'existence d'une limite à $m(N)$ lorsque $N \to \infty$.

Pour terminer essayons de décrire (de façon très conjoncturelle) ce qui semble se produire dans le modèle à p spins, par rapport à la situation du REM. Dans ce dernier, la fonction d'énergie $H_N(\sigma)$ est "totalement discontinue" et ses plus petites valeurs ont asymptotiquement une distribution que nous avons étudiée. Dans le modèle à p-spins la fonction d'énergie est maintenant "continue" mais elle présente néanmoins des minimums locaux. Les plus petits d'entre eux sont situés sur des configurations essentiellement orthogonales et leurs valeurs, en un certain sens, s'organisent comme dans le cas du REM (et les "états" se forment dans leurs voisinages).

15. CONCLUSION

Ce bref exposé a tenté de donner un aperçu de quelques progrès récents, mais aussi de montrer la variété et le caractère naturel, sinon fondamental, de plusieurs questions sur lesquelles tout ou presque reste à faire, et qui, espérons-le, attireront dans le futur les efforts de nombreux chercheurs.

Il a fallu, faute de place, limiter le sujet à "la statique du champ moyen", et en particulier il n'a pas été possible d'aborder les remarquables travaux récents de Guionnet et Ben Arous sur la "dynamique", qui constitue un sujet relié mais très différent de celui traité ici, et pour lequel le lecteur devra se référer aux articles de ces auteurs [G], [BA-G].

BIBLIOGRAPHIE

[A-L-R] M. AIZENMAN, J.L. LEBOWITZ, D. RUELLE - *Some rigorous results on the Sherrington-Kirkpatrick model*, Commun Math. Phys. **112** (1987), 3-20.

[A-G-S] D.J. AMIT, H. GUTFREUND, H. SOMPOLINSKY - *Statistical Mechanics of Neural Networks near Saturation*, Annals of Physics **173** (1987), 30-67.

[BA-G] G. BEN AROUS, A. GUIONNET - *Langevin spin glass dynamic*, A. Bovier et P. Picco editors, Progress in Probability, vol. **41**, Birkhäuser, Boston (1998).

[B-S] E. BOLTHAUSEN, A.S. SZNITMAN - *On Ruelle's probability casades and an astract cavity method*, Comm. Math. Phys. **197** (1997), 247-276.

[B-G] A. BOVIER, V. GAYRARD - *Hopfield models as generalized random mean field models*, Mathematical aspects of spin glasses and neural networks, A. Bovier and P. Picco editors, Progress in Probability, vol **41**, Birkhäuser, Boston (1997).

[B-G-P] A. BOVIER, V. GAYRARD, P. PICCO - *Gibbs states of the Hopfield model with extensively many patterns*, J. Stat. Phys. **79** (1995), 395-414.

[B-P] A. BOVIER, P. PICCO (editors), - *Mathematical aspects of Spin Glasses and Neural Networks*, Progress in Probability, vol. **41**, Birkhäuser, Boston (1997).

[C] F. COMETS - *A spherical bound for the Sherrington-Kirkpatrick model*, in "Hommage à P.-A. Meyer et J. Neveu", Astérisque **236** (1996), 103-108.

[C-N] F. COMETS, J. NEVEU - *Sherrington-Kirkpatrick model of spin glasses and stochastic calculus : the high temperature case*, Commun Math. Phys. **166** (1995), 549-564.

[D] B. DERRIDA - *Random energy model : An exactly solvable model of disordered systems*, Phy. Rev. B, **24** ♯ 5 (1981), 2613-2626.

[D-G] B. DERRIDA, E. GARDNER - *Optimal storage properties of Neural Network models*, J. Phy. A. **21** (1988), 271-284.

[F-Z1] J. FRÖHLICH, B. ZEGARLINSKI - *The high-temperature phase of long-range spin glasses*, Comm. Math. Phys. **110** (1987), 121-155.

[F-Z2] J. FRÖHLICH, B. ZEGARLINSKI - *Some comments on the Sherrington-Kirkpatrick model of spin glasses*, Comm. Math. Phy. **112** (1987), 553-566.

[F] F. FERNIQUE - *Régularité des trajectoires des fonctions aléatoires gaussiennes*, École d'été des Probabilités de Saint-Flour IV, Lecture Notes in Math., **480** (1975), 1-96, Springer-Verlag.

[G1] E. GARDNER - *Spin glasses with p-spin interactions*, Nuclear Phys. B **257**, ♯ 6 (1985), 747-765.

[G2] E. GARDNER - *The space of interactions in neural network models*, J. Phys. A. **21** (1988), 257-270.

[G] A. GUIONNET - *Averaged and quenched propagation of chaos for spin glass dynamics*, Probab. Theor. relat. fields **109** (1997), 183-215.

[H] J.J. HOPFIELD - *Neural networks and physical systems with emergent collective computational abilities*, Proc. Natl. Acad. Sci. USA **79** (1982), 1554-2558.

[H-K-P] J. HERTZ, A. KROGH, R.C. PALMER - *Introduction to the theory of Neural computation*, Addison Wesley (1991).

[K-T-J] J.M. KOSTERLITZ, D.J. THOULESS, K.C. JONES - *Spherical model of a Spin-Glass*, Phys. Rev. letters **36** (1976), 1217-1220.

[Lou] D. LOUKIANOVA - *Lower bounds on the restitution error of the Hopfield model*, Probab. Th. Relat. Fields **107** (1997), 161-176.

[M] M. MEZARD - *The space of interactions in neural networks : Gardner's computation with the cavity method*, J. Phys. A. **22** (1988), 2281-2190.

[M-P-V] M. MEZARD, G. PARISI, M. VIRASIRO - *Spin glass theory and beyond*, World Scientific, Singapore (1987).

[M-Z] R. MONASSON, R. ZECCHINA - *Statistical Mechanics of the Random K-Sat Model*, Phys. Rev. E (1997), 1357-1370.

[N] C. NEWMAN - *Memory capacity in neural network models : Rigorous lower bounds*, Neural Networks, **I** (1988), 223-238.

[N-S] C. NEWMAN, D. STEIN - *Thermodynamic chaos and the structure of short-range spin glasses, Mathematical aspects of Spin Glass and Neural networks*, A. Bovier and P. Picco editors, Progress in Probability, vol. **41**, Birkhäuser, Boston (1998), 243-287.

[P-S] L. PASTUR, M. SHCHERBINA - *Absence of Self-Averaging of the order parameter in the Sherrington-Kirkpatrick model*, J. Stat. Phys. **62** (1991), 1-19.

[P-Y] J. PITMAN, M. YOR - *The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator*, Ann. Probab. **25** (1997), 855-900.

[R] D. RUELLE - *A mathematical reformulation of Derrida's REM and GREM*, Comm. Math. Phys. **108** (1987), 225-239.

[Sh1] M. SHCHERBINA -*More about absence of self averaging of the order parameter in the Sherrington-Kirkpatrick model*, CARR Reports in Mathematical Physics, n° **3/91**, Department of Mathematics, University of Rome "la Sapienza", (1991).

[Sh2] M. SHCHERBINA - *On the Replica-Symmetric Solution for the Sherrington-Kirkpatrick Model*, Helv. Phys. Acta **70** (1997), 838-853.

[S-K] D. SHERRINGTON, S. KIRKPATRICK - *Solvable model of a spin glass*, Phys. Rev. Lett. **35** (1972), 1792-1796.

[T1] M. TALAGRAND - *Regularity of Gaussian processes*, Acta Math. **159** (1987), 99-149.

[T2] M. TALAGRAND - *Concentration of measure and isoperimetric inequalities in product spaces*, Publ. Math. I.H.E.S. **81** (1995), 73-205.

[T3] M. TALAGRAND - *The Sherrington-Kirkpatrick model: A challenge for mathematicians*, Probab. Theor. Relat. Fields **110** (1998), 109-176.

[T4] M. TALAGRAND - *Rigorous results for the Hopfield model with many patterns*, Probab. Theory Relat. Fields **110** (1998), 177-276.

[T5] M. TALAGRAND - *Huge random structures and mean field models for spin glasses*, Proceedings of the Berlin international congress of mathematicians, Documenta Math. (1998).

[T6] M. TALAGRAND - *La page de Michel Talagrand*, http : // www.proba.jussieu.fr.

Michel TALAGRAND

Université de Paris 6
URA 7064 du CNRS
4 place Jussieu
F–75252 PARIS CEDEX

E–mail : mit@ccr.jussieu.fr

Astérisque

CHRISTOPHE BREUIL

Intégration sur les variétés p-adiques

Astérisque, tome 266 (2000), Séminaire Bourbaki,
exp. n° 860, p. 319-350

<http://www.numdam.org/item?id=SB_1998-1999__41__319_0>

© Société mathématique de France, 2000, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (http://smf4.emath.fr/Publications/Asterisque/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Séminaire BOURBAKI
51ème année, 1998-99, n° 860

Juin 1999

INTÉGRATION SUR LES VARIÉTÉS p-ADIQUES
[d'après Coleman, Colmez]

par **Christophe BREUIL**

TABLE DES MATIÈRES

Introduction 320
Conventions et notations 321
1. Construction de l'intégration p-adique 322
1.1. La construction de Colmez 322
1.1.1. Énoncé du théorème 322
1.1.2. Description de la preuve 323
1.1.3. Le point de vue de Zarhin 327
1.2. La construction de Coleman 327
1.2.1. Quelques mots d'introduction sur la géométrie rigide 328
1.2.2. Intégration des formes de seconde espèce 329
2. Applications 333
2.1. L'accouplement des périodes p-adiques 333
2.1.1. Quelques rappels sur B_{dR}^{+} 334
2.1.2. Accouplement des périodes p-adiques : énoncé du théorème 335
2.1.3. Esquisse de preuve et relations de Riemann p-adiques 337
2.2. Les polylogarithmes p-adiques 340
2.2.1. Itération de l'intégration sur les courbes 340
2.2.2. Construction des polylogarithmes p-adiques 343
2.3. Autres applications 346
RÉFÉRENCES 347

Introduction

Soit X une variété algébrique lisse sur le corps $\mathbf{R}$ des nombres réels. Les points de X à valeurs dans $\mathbf{R}$ ou dans $\mathbf{C}$, que l'on note $X(\mathbf{R})$ et $X(\mathbf{C})$, forment une variété analytique réelle ou complexe, sur laquelle on dispose de tout l'arsenal de la théorie classique de l'intégration, développé principalement avant le début du XX^e siècle. Depuis l'introduction par Hensel des nombres p-adiques et de la topologie p-adique dans les années 1900, les mathématiciens n'ont eu de cesse d'étendre à $\mathbf{Q}_p$ ce qu'ils savaient faire sur $\mathbf{R}$. Ainsi peut-on se demander s'il existe une théorie de l'intégration sur $X(\mathbf{Q}_p)$ lorsque X est une variété algébrique lisse sur $\mathbf{Q}_p$. Or, la topologie p-adique est très différente de la topologie classique car elle est totalement discontinue, c'est-à-dire que tout point possède une base de voisinages à la fois ouverts et fermés ou encore qu'il n'y a pas d'obstruction à passer du local au global ! Comme tout le charme réside essentiellement dans cette obstruction, cette tâche a longtemps semblé vouée à des sorites sans grand intérêt. Pourtant, l'existence du logarithme p-adique, connue dès les tout débuts, fournissait un exemple tout à fait non trivial d'une fonction p-adique qui avait bien l'air d'une "primitive p-adique"...

La situation change au début des années soixante avec l'introduction par Tate des espaces analytiques rigides ([Ta2]). L'idée en soi est naturelle : puisque la topologie p-adique a trop d'ouverts, on va en sélectionner certains plus "jolis" que les autres. On connaît depuis l'essor de cette théorie, dont l'un des aboutissements fut les théorèmes d'uniformisation p-adique de Mumford et Raynaud ([Mu2], [Ra3]). Coleman le premier, dans les années 80, réussit à utiliser la "topologie rigide" pour définir dans [Cole3] une théorie de l'intégration des 1-formes différentielles algébriques de seconde espèce sur les variétés algébriques projectives (ou propres) et lisses sur $\mathbf{Q}_p$ avec bonne réduction. Motivé par les applications aux fonctions L p-adiques, il étend dans le cas des courbes son intégration à toutes les 1-formes rationnelles, et même à des formes non nécessairement algébriques ou rigides, ce qui lui permet par exemple de définir des polylogarithmes p-adiques. Enfin, au début des années 90, Colmez dans [Colm1], et indépendamment Zarhin dans [Za1], parviennent à rendre possible l'impossible en intégrant sur du totalement discontinu : non seulement ils étendent l'intégrale de Coleman au cas de variétés algébriques supposées seulement lisses sur $\mathbf{Q}_p$ et à toutes les 1-formes rationnelles fermées, mais ils le font sans passer par la géométrie rigide. Cependant, leurs méthodes ne s'appliquent pour l'instant qu'aux 1-formes algébriques et ne permettent pas de retrouver les polylogarithmes p-adiques par exemple. Une fois débarrassé des hypothèses de propreté et de bonne réduction, Colmez généralise alors de façon systématique dans [Colm2] beaucoup des théorèmes de Coleman, et obtient en outre une preuve originale et particulièrement élégante, car plus proche de son analogue complexe, de l'accouplement des périodes p-adiques des variétés abéliennes. Dernièrement, Besser a donné dans [Bes] une généralisation de l'intégrale de Coleman (sur les variétés propres lisses avec bonne réduction) au cas de n-formes fermées avec $n > 1$.

L'objectif de cet exposé est de présenter dans une première partie l'intégration de Colmez et l'intégration de Coleman, puis dans une seconde partie certaines de leurs applications. Il a été conçu plutôt comme une introduction au sujet et non comme un "survey". Ainsi, j'ai pris le parti d'exposer de façon assez complète les deux constructions de l'intégration, puis de développer avec quelques détails deux applications particulières (essentiellement une par auteur), enfin de mentionner très brièvement quelques autres. L'exposé n'est donc pas exhaustif (certains aspects ne sont pas du tout abordés) et je prie le lecteur - ou l'auteur - averti de me pardonner s'il ne trouve pas développée ici son application favorite.

Durant la préparation de cet exposé, j'ai bénéficié de conversations orales ou "électroniques" avec plusieurs personnes dont : Amnon Besser, Jean-Benoît Bost, Robert Coleman, Pierre Colmez, Jean-Louis Colliot-Thélène, Philippe Gille, Bernard Le Stum, William Messing et Michel Raynaud. Qu'ils soient tous chaleureusement remerciés pour leur aide. Je suis reconnaissant à Amnon Besser, Robert Coleman et Glenn Stevens de m'avoir envoyé respectivement les références [Bes], [CI] et [St] (parmi d'autres ; la lecture de [St] en particulier m'a facilité la rédaction de la partie 2.2). Enfin, je remercie Jean-Benoît Bost, Laurent Clozel, Jean-Marc Fontaine, Michel Raynaud et surtout Pierre Colmez pour leurs remarques sur des versions préliminaires de ce texte.

Conventions et notations

Dans tout cet article, on fixe un nombre premier p et une clôture algébrique $\overline{\mathbf{Q}}_p$ de $\mathbf{Q}_p$. On note $\mathbf{C}_p$ la complétion p-adique de $\overline{\mathbf{Q}}_p$, $\mathcal{O}_{\mathbf{C}_p}$ l'anneau des entiers de $\mathbf{C}_p$ et $|\cdot|$ la valeur absolue p-adique sur $\mathbf{C}_p$ donnée par $|x| = \frac{1}{p^{\mathrm{val}(x)}}$ où val est la valuation p-adique normalisée par $\mathrm{val}(p) = 1$. Le logarithme d'Iwasawa est l'unique fonction $\log : \mathbf{C}_p - \{0\} \to \mathbf{C}_p$ telle que :

$$\begin{aligned} \log(1-x) &= -\sum_{n=1}^{+\infty} \frac{x^n}{n} \text{ si } |x| < 1 \\ \log(xy) &= \log(x) + \log(y) \\ \log(p) &= 0. \end{aligned}$$

Le choix $\log(p) = 0$ est arbitraire (on peut même décider que $\log(p)$ est une indéterminée, cf. [Colm2]) mais naturel du point de vue de la théorie des nombres. Une variété est un schéma X séparé de type fini et géométriquement intègre sur un corps K. Nous ne considérerons en fait que des variétés lisses, c'est-à-dire telles que $X \to \mathrm{Spec}(K)$ est un morphisme lisse. Une variété abélienne sur K est une variété X sur K munie d'une structure de K-schéma en groupes et telle que $X \to \mathrm{Spec}(K)$ est propre. Cela entraîne $X \to \mathrm{Spec}(K)$ lisse. Une fonction rationnelle (resp. une 1-forme rationnelle) sur une

variété X est un élément du corps des fonctions $K(X)$ (resp. des différentielles de Kähler $\Omega^1_{K(X)/K}$) qu'on peut voir comme une fonction (resp. une 1-forme) régulière sur un ouvert de Zariski U de X. Si X est une variété sur K, on supposera toujours $X(K) \neq \emptyset$.

1. CONSTRUCTION DE L'INTÉGRATION p-ADIQUE

1.1. La construction de Colmez

Dans cette partie, K désigne soit $\mathbf{C}_p$, soit une extension finie de $\mathbf{Q}_p$ dans $\mathbf{C}_p$.

1.1.1. *Énoncé du théorème.* Soit X une variété lisse sur K de dimension d. On a supposé $X(K) \neq \emptyset$. Pour tout ouvert de Zariski affine U de X, il existe une immersion fermée $U \hookrightarrow \mathbf{A}^m$ pour $m \gg 0$ et on définit la topologie p-adique sur $U(K)$ comme la topologie induite par la topologie p-adique sur $\mathbf{A}^m(K) = K^m$. C'est indépendant de l'immersion choisie. Par recollement, on obtient la topologie p-adique sur $X(K)$. On appelle ouverts p-adiques les ouverts pour cette topologie. Tout point de $X(K)$ admet pour la topologie p-adique une base de voisinages ouverts et fermés puisque tel est le cas sur K^m : cette topologie est donc totalement discontinue. Comme X est lisse, pour tout $x \in X(K)$, il existe un ouvert de Zariski $U_x = \mathrm{Spec}(K[x_1, ..., x_n, z_1, ..., z_d]/(f_1, ..., f_n))$ contenant x où $(\frac{\partial f_i}{\partial x_j})_{1 \leq i,j \leq n}$ est de déterminant inversible et tel que $x \in U_x(K)$ correspond à $(x_i = 0, z_j = 0)$. Pour définir la structure *analytique* p-adique sur $X(K)$, rappelons le :

Lemme 1.1.1.1. — *La projection sur les d dernières composantes $z_1, ..., z_d$ induit un homéomorphisme entre un ouvert p-adique V de $U_x(K) \subset X(K)$ contenant x et la boule ouverte $B(0, \delta_-)^d = \{(z_1, ..., z_d) \in K^d \mid |z_i| < \delta,\ 1 \leq i \leq d\}$ pour un δ suffisamment petit dans $\mathbf{R}^+ - \{0\}$.*

Cela résulte d'un algorithme élémentaire (théorème des fonctions implicites !), voir par exemple ([Colm2],A.3.4). Remarquons que tout ouvert p-adique peut s'écrire comme une réunion *disjointe* de boules ouvertes comme en (1.1.1.1) (car si B_1^d et B_2^d sont deux telles boules, on a soit $B_1^d \cap B_2^d = \emptyset$, soit $B_1^d \subset B_2^d$, soit $B_2^d \subset B_1^d$).

Définition 1.1.1.2. — *Une fonction localement analytique sur un ouvert p-adique W de $X(K)$ est une fonction f de W dans K telle que, pour tout $x \in W$ et pour un choix (donc tout choix) de paramètres locaux z_i autour de x comme en (1.1.1.1), il existe $S \in K[[z_1, ..., z_d]]$ et $r \in \mathbf{R}^+ - \{0\}$ ($r \leq \delta$) tels que $S(z_1, ..., z_d)$ converge si $|z_i| < r$ et coïncide avec f sur ce voisinage.*

Exemple 1.1.1.3. — Toute fonction rationnelle f régulière sur un ouvert de Zariski U est une fonction localement analytique sur $U(K)$. Soit $\log(f)$ la fonction définie comme la composée $U(K) - \{x \mid f(x) = 0\} \xrightarrow{f} K^\times \xrightarrow{\log} K$: $\log(f)$ est une fonction localement analytique sur $W = U(K) - \{x \mid f(x) = 0\}$.

On définit l'espace analytique p-adique X^{an} associé à X comme l'espace annelé $X(K)$ muni de la topologie p-adique et du faisceau $\mathcal{O}_{X^{an}}$ des fonctions localement analytiques. Si $\mathcal{O}_X$ est le faisceau structural sur X, $\mathcal{O}_{X^{an}}$ restreint à la topologie de Zariski est un faisceau de $\mathcal{O}_X$-modules et, en utilisant (1.1.1.1), on peut définir une différentielle d : $\mathcal{O}_{X^{an}} \to \mathcal{O}_{X^{an}} \otimes_{\mathcal{O}_X} \Omega^1_{X/K}$ où $\Omega^1_{X/K}$ est le faisceau des différentielles de Kähler. Si f est une fonction rationnelle, on a $d\log(f) = \frac{df}{f}$. On note $K(X)_{log} = K(X)[\log(f),\ f \in K(X)^\times]$ où les éléments de $K(X)$ sont vus comme des fonctions localement analytiques. Les fonctions f de $K(X)_{log}$ sont toutes localement analytiques sur un ouvert de Zariski $U_f(K)$ de $X(K)$. Le lemme suivant sera utilisé plusieurs fois :

Lemme 1.1.1.4. — *Soient $f \in K(X)_{log}$ et $W \subset X(K)$ un ouvert p-adique. Si $f|_W = 0$ alors $f = 0$.*

Voir ([Colm2],II.1.7) pour la preuve. Inutile de dire que cet énoncé est faux pour des fonctions localement analytiques quelconques.

Notre objectif est de définir une théorie de l'intégration pour les 1-formes rationnelles fermées sur X. À cause du lemme de Poincaré, toute 1-forme fermée s'intègre localement pour la topologie p-adique sur une boule suffisamment petite, donc s'intègre globalement. Le problème est qu'il y a une multitude de constantes locales d'intégration : une sur chaque boule ! Et pourtant, parmi cette infinité de primitives, il en existe miraculeusement une plus jolie que les autres :

Théorème 1.1.1.5. — ([Colm2], [Za2]) *Il existe une unique façon d'associer à une variété X lisse sur K et une 1-forme rationnelle fermée ω sur X une fonction f_ω bien définie à addition près d'une constante et localement analytique sur $X(K) \setminus \{\text{pôles de } \omega\}$, de sorte que :*

(i) $df_\omega = \omega$

(ii) $f_{\lambda\omega+\lambda'\omega'} = \lambda f_\omega + \lambda' f_{\omega'}$ $(\lambda, \lambda' \in K)$

(iii) $f_\omega = f$ *si* $\omega = df$ *où* $f \in K(X)_{log}$

(iv) *si* $g : X' \to X$ *est un morphisme de variétés lisses sur* $\mathrm{Spec}(K)$, $f_{g^*\omega} = g^* f_\omega$ $(= f_\omega \circ g)$.

1.1.2. *Description de la preuve.* Nous donnons ici la preuve de Colmez (voir 1.1.3 pour Zarhin). Elle se coupe en deux : 1) en utilisant les morphismes d'Albanese on montre qu'il suffit de définir une intégration sur les variétés abéliennes, 2) on construit effectivement cette intégration sur les variétés abéliennes.

Commençons par 1). Rappelons qu'à toute variété X lisse sur K, on peut associer de manière fonctorielle sa variété d'Albanese $Alb(X)$ (= une variété abélienne) encore définie sur K (lorsque X est propre et lisse sur $\mathbf{C}$, $Alb(X)$ est isomorphe en tant que tore complexe à $H^1_{dR}(X)^*/H_1(X(\mathbf{C}), \mathbf{Z})$). Tout point $P \in X(K)$ donne lieu à un morphisme

unique $\iota_P : X \to Alb(X)$ défini sur K tel que $\iota_P(P) = 0$. Un tel morphisme s'appelle un morphisme d'Albanese (pour tout cela, voir [Se1] ou [La] ou ([Colm2],I.5)). De plus, si X est propre, ι_P induit un isomorphisme $H^1_{dR}(Alb(X)) \tilde{\to} H^1_{dR}(X)$.

Lemme 1.1.2.1. — *Soient $\iota_P : X \to Alb(X)$ un morphisme d'Albanese et ω une 1-forme rationnelle fermée sur X, alors il existe une 1-forme rationnelle fermée η sur $Alb(X)$, des fonctions rationnelles $f, f_1, ..., f_n$ sur X et des constantes $\lambda_1, ..., \lambda_n$ dans K telles que :*

$$\omega = \iota_P^* \eta + df + \sum_{i=1}^{n} \lambda_i \frac{df_i}{f_i}.$$

Indication de preuve. — Soit $\tilde{H}^1(X)$ (resp. $\tilde{H}^1(Alb(X))$ le quotient des 1-formes rationnelles fermées par $dK(X)_{log}$ (resp. $dK(Alb(X))_{log}$), il suffit de montrer que ι_P induit un isomorphisme $\tilde{H}^1(Alb(X)) \tilde{\to} \tilde{H}^1(X)$ et il suffit pour cela d'étendre les scalaires à une clôture algébrique $\overline{K}$ de K. En remarquant qu'alors, si $\overline{X}$ est une compactification lisse de X (qui existe par la résolution des singularités d'Hironaka), on a $\tilde{H}^1(X) \simeq \tilde{H}^1(\overline{X})$ (car $\tilde{H}^1(X)$ ne dépend que du corps des fractions de X) et $Alb(X) \simeq Alb(\overline{X})$, on est ramené au cas d'une variété X propre et lisse sur $\overline{K}$. On a dans ce cas un diagramme commutatif de suites exactes avec des isomorphismes à gauche et à droite ([Colm2],I.1.16) :

$$\begin{array}{ccccccccc} 0 & \to & H^1_{dR}(X) & \to & \tilde{H}^1(X) & \to & \overline{K} \otimes_{\mathbf{Z}} Pic^0(X) & \to & 0 \\ & & \wr\uparrow & & \uparrow & & \uparrow\wr & & \\ 0 & \to & H^1_{dR}(Alb(X)) & \to & \tilde{H}^1(Alb(X)) & \to & \overline{K} \otimes_{\mathbf{Z}} Pic^0(Alb(X)) & \to & 0 \end{array}$$

où les flèches horizontales de droite sont les applications qui à une forme différentielle associe son résidu, un diviseur algébriquement équivalent à 0 (voir [BLR] pour Pic^0). On en déduit le résultat. □

Proposition 1.1.2.2. — *Supposons que l'énoncé* (1.1.1.5) *soit vrai pour les variétés abéliennes sur K (i.e. en ne considérant en* (iv) *que des morphismes entre variétés abéliennes), alors il est vrai pour toutes les variétés algébriques lisses sur K.*

Preuve. — Pour toute 1-forme rationnelle fermée η sur une variété abélienne, on a donc une primitive f_η bien définie à une constante près. Soit ω une 1-forme rationnelle fermée sur une variété X lisse sur K, $\iota_P : X \to Alb(X)$ un morphisme d'Albanese et η, f, f_i, λ_i comme en (1.1.2.1). Si f_ω comme en (1.1.1.5) existe, alors $f_\omega = \iota_p^* f_\eta + f + \sum \lambda_i \log(f_i)$. Si $\omega = \iota_P^* \eta' + df' + \sum \lambda'_i \frac{df'_i}{f'_i}$ est une autre écriture, on a $\iota_P^*(\eta - \eta') \in dK(X)_{log}$, d'où $\eta - \eta' \in dK(Alb(X))_{log}$ puisque $\tilde{H}^1(X) \simeq \tilde{H}^1(Alb(X))$ (cf. preuve de 1.1.2.1), et $f_\eta - f_{\eta'} \in K(Alb(X))_{log}$ d'après la propriété (iii). Soit $f'_\omega = \iota_p^* f_{\eta'} + f' + \sum \lambda'_i \log(f'_i) : f_\omega - f_{\omega'} \in K(X)_{log}$ et $d(f_\omega - f_{\omega'}) = 0$, donc $f_\omega - f_{\omega'}$ est constante sur un ouvert p-adique, donc partout (1.1.1.4). La primitive f_ω modulo les constantes est donc indépendante des choix. Quitte à changer les écritures de ω, on vérifie qu'elle s'étend bien en une fonction localement analytique sur $X(K) \setminus \{\text{pôles de } \omega\}$. Il reste à montrer la propriété (iv), mais si $g : X' \to X$ est un morphisme défini sur

K, $P \in X(K)$, $Q \in X'(K)$, il existe un (unique) morphisme $Alb(g) : Alb(X') \to Alb(X)$ tel que le diagramme :

$$\begin{array}{ccc} X' & \xrightarrow{g} & X \\ {\scriptstyle \iota_P}\downarrow & & \downarrow{\scriptstyle \iota_Q} \\ Alb(X') & \xrightarrow{Alb(g)} & Alb(X) \end{array}$$

soit commutatif (cela résulte de la propriété universelle de la variété d'Albanese [Se1]). La propriété (iv) découle facilement de la construction de f_ω ci-dessus et du fait que $f_\eta \circ Alb(g) = f_{Alb(g)^*\eta}$ si η est une 1-forme rationnelle fermée sur $Alb(X)$. □

Il faut maintenant définir ces primitives sur les variétés abéliennes. Si X est une variété abélienne, on note, pour $I \subset \{1,2\}$, $m_I : X^3 \to X$ l'application qui à $(x, h_1, h_2) \in X^3$ associe $x \oplus (\oplus_{i\in I} h_i)$ où $\oplus$ est la loi d'addition sur X (plus exactement sur le faisceau en groupes représenté). Si α est une fonction, ou une 1-forme, rationnelle sur X, on pose $\Delta^{[2]}\alpha = m^*_{\{1,2\}}\alpha - m^*_{\{1\}}\alpha - m^*_{\{2\}}\alpha + m^*_\emptyset\alpha$. Les deux résultats suivants vont jouer un rôle clef :

Théorème 1.1.2.3. — *Soit ω une 1-forme rationnelle fermée sur X, alors $\Delta^{[2]}\omega \in dK(X^3)_{log}$.*

Théorème 1.1.2.4. — (i) *Les sous-groupes ouverts de $X(K)$ forment un système fondamental de voisinages de l'origine pour la topologie p-adique.*

(ii) *Si V est un sous-groupe ouvert de $X(K)$, alors $X(K)/V$ est un groupe de torsion.*

Le théorème (1.1.2.3) est une des incarnations du théorème du carré ([Mu1],6 cor.4). C'est un énoncé algébrique qui n'a rien de spécifique à la topologie p-adique. Pour une preuve sur $\overline{K}$ sous cette forme via le principe de Lefschetz, voir ([Colm2],I.3.6), le résultat sur K s'en déduisant en prenant les invariants sous $Gal(\overline{K}/K)$. Le théorème (1.1.2.4) est lui par contre tout à fait particulier à la topologie p-adique. Un des arguments clef de sa preuve est que les boules ouvertes comme en (1.1.1.1) sont des groupes pour l'addition sur K^d, ce qui est rarement le cas des boules ouvertes de $\mathbf{C}^d$! Voir ([Colm2],II.1.9) ou ([Cole4],4.1) pour des preuves de (1.1.2.4).

Si une primitive f_ω de ω satisfaisant les conditions du théorème existe, on doit avoir $f_{\Delta^{[2]}\omega} \in K(X^3)_{log}$ et $\Delta^{[2]}f_\omega = f_{\Delta^{[2]}\omega}$ d'après les conditions (iii) et (iv) en (1.1.1.5), ce qui entraîne $\Delta^{[2]}f_\omega \in K(X^3)_{log}$.

Théorème 1.1.2.5. — *Soient ω une 1-forme rationnelle fermée sur X et U_ω l'ouvert de Zariski de X complémentaire des pôles de ω, il existe à addition près d'une constante une unique fonction f_ω localement analytique sur $U_\omega(K)$ telle que :*

1) $df_\omega = \omega$

2) $\Delta^{[2]}f_\omega \in K(X^3)_{log}$.

Preuve. — En utilisant $\Delta^{[2]}T_a^* = T^*_{(a,0,0)}\Delta^{[2]}$ (où $a \in X(K)$ et T_a est la translation par a sur $X(K)$), on se ramène au cas où ω est définie en 0.

Unicité : S'il y a deux primitives, leur différence D_ω satisfait $dD_\omega = 0$, donc D_ω est localement constant et on a aussi $d(\Delta^{[2]}D_\omega) = 0$. Mais $\Delta^{[2]}D_\omega \in K(X^3)_{log}$, d'où $\Delta^{[2]}D_\omega = 0$ par (1.1.1.4) car $\Delta^{[2]}D_\omega|_{X\times\{0\}\times\{0\}} = 0$. Pour $h_1, h_2 \in U_\omega(K)$ tels que $h_1 \oplus h_2 \in U_\omega(K)$, on a donc, quitte à remplacer D_ω par $D_\omega - D_\omega(0)$, $D_\omega(h_1 \oplus h_2) = D_\omega(h_1) + D_\omega(h_2)$. Cela entraîne que D_ω se prolonge par continuité en un morphisme de groupes sur tout $X(K)$: si V est un sous-groupe ouvert de $U_\omega(K)$ sur lequel D_ω est défini et $x \in X(K)$, d'après (1.1.2.4,ii) il existe $n_x \in \mathbf{N}$ tel que $[n_x]x = x \oplus ... \oplus x$ (n_x fois) est dans V et on pose $D_\omega(x) = \frac{1}{n_x}D_\omega([n_x]x)$. Comme D_ω est localement constant, on peut choisir un tel V sur lequel D_ω est constant. Mais $D_\omega|_V$ est alors nul puisque c'est un homomorphisme de groupes, donc D_ω est nul partout par la formule ci-dessus.

Existence : D'après (1.1.2.3), il existe $g \in K(X^3)_{log}$ tel que $\Delta^{[2]}\omega = dg$ et on peut supposer g nulle sur $X(K) \times \{0\} \times \{0\}$ puisque $\Delta^{[2]}\omega|_{X\times\{0\}\times\{0\}} = 0$.
Première étape : Soit V un sous-groupe ouvert de $X(K)$ contenu dans $U_\omega(K)$ sur lequel ω est analytique. On peut prendre V isomorphe à une boule ouverte $B(0, \delta_-)^d$ comme en (1.1.1.1). Soit f_ω l'unique primitive analytique de ω sur V nulle en 0. On a $d(\Delta^{[2]}f_\omega - g|_{V^3}) = \Delta^{[2]}\omega - \Delta^{[2]}\omega = 0$ donc $\Delta^{[2]}f_\omega = g|_{V^3}$ car ces deux fonctions sont analytiques sur V^3 et nulles sur $V \times \{0\} \times \{0\}$.
Deuxième étape : On prolonge f_ω à $U_\omega(K)$. Notons V_n l'ouvert p-adique de $X(K)$ formé des x tels que $[n]x \in V$ et $[k]x$ n'est pas un pôle de ω pour tout $k \in \{1, ..., n\}$. On étend f_ω à V_n en posant :

$$f_\omega(x) = \frac{1}{n}\Big(f_\omega([n]x) - \sum_{k=1}^{n-1} g(0, [k]x, x)\Big).$$

En utilisant $f_\omega(x \oplus y) - f_\omega(x) - f_\omega(y) = g(0, x, y)$ sur V, on vérifie que cette équation est en particulier satisfaite sur V. De plus, $f_\omega(x)$ est indépendant de n tel que $x \in V_n$ et on a $\Delta^{[2]}f_\omega = g$ là où les deux termes sont définis. En effet, ces assertions se ramènent facilement à des égalités fonctionnelles sur g seulement, qui sont automatiquement satisfaites sur V et on peut utiliser (1.1.1.4) puisque $g \in K(X^3)_{log}$ pour en déduire qu'elles sont vraies partout. On obtient une fonction f_ω localement analytique sur $\cup_n V_n$ qui est un ouvert p-adique dense de $X(K)$ contenu dans $U_\omega(K)$ et telle que $\Delta^{[2]}f_\omega = g$ partout où ça a un sens. Quitte à translater 0 en un point $a \in U_\omega(K)$ et à utiliser $T^*_{(\ominus a,0,0)}\Delta^{[2]}T_a^* = \Delta^{[2]}$, le même raisonnement appliqué à $T_a^*\omega$ fournit une autre fonction $f_{\omega,a}$ localement analytique sur un ouvert p-adique dense contenant 0 et telle que $\Delta^{[2]}(f_\omega - T^*_{\ominus a}f_{\omega,a}) = 0$. Soit b un point où les deux fonctions f_ω et $T^*_{\ominus a}f_{\omega,a}$ sont définies, la fonction $h(x) = (f_\omega - T^*_{\ominus a}f_{\omega,a})(b \oplus x) - T^*_{\ominus a}f_{\omega,a}(b) + f_\omega(b)$ satisfait $h(x \oplus y) = h(x) + h(y)$ partout où elle est définie, et cela entraîne qu'elle se prolonge par continuité à $X(K)$ tout

entier (cf. la partie sur l'unicité). Donc f_ω se prolonge par continuité en a, c'est-à-dire finalement à tout $U_\omega(K)$.
Troisième étape : Soit f_ω comme dans la deuxième étape et $\eta = df_\omega - \omega$, on a $\Delta^{[2]}\eta = 0$, d'où $(\Delta^{[2]}\eta)|_{\{0\}\times\{a\}\times X} = 0$ i.e. $T_a^*\eta = \eta$. Donc η est une 1-forme localement analytique invariante par translation et nulle sur V, donc nulle partout. □

Si P et Q sont deux points de $X(K)$ où ω est définie, posons $\int_P^Q \omega = f_\omega(Q) - f_\omega(P)$.

Corollaire 1.1.2.6. — *Avec les notations précédentes* :
(i) $\int_P^Q(\lambda\omega + \lambda'\omega') = \lambda\int_P^Q \omega + \lambda'\int_P^Q \omega'$ $(\lambda, \lambda' \in K)$
(ii) $\int_P^Q \omega = f(Q) - f(P)$ *si* $\omega = df$ *où* $f \in K(X)_{log}$
(iii) *si* $g : X' \to X$ *est un morphisme de variétés abéliennes sur* $\mathrm{Spec}(K)$, $\int_P^Q g^*\omega = \int_{g(P)}^{g(Q)} \omega$.

Preuve. — Les énoncés (i) à (iii) résultent de l'unicité modulo cte de f_ω et de $g^*\Delta^{[2]} = \Delta^{[2]}g^*$. □

Avec (1.1.2.2), cela achève la preuve du théorème (1.1.1.5).

Remarque 1.1.2.7. — On peut préciser le comportement des primitives f_ω au voisinage des pôles de ω : c'est la somme d'une fonction localement méromorphe et d'une combinaison linéaire de logarithmes de fonctions localement méromorphes.

Remarque 1.1.2.8. — Si X est une variété abélienne et $\omega \in H^0(X, \Omega^1_{X/K})$, ω est une 1-forme fermée invariante par translation sur X ([Mu1],II.4.iii) et f_ω est alors un homomorphisme de groupes de $X(K)$ dans K. Selon la terminologie de ([Bou],III.7), une telle primitive est appelée un *logarithme de* X.

1.1.3. *Le point de vue de Zarhin.* Dans [Za1] et [Za2], Zarhin a construit indépendamment des primitives à toutes les 1-formes rationnelles fermées sur les variétés lisses qui satisfont toutes les propriétés en (1.1.1.5), et donc coïncident avec les primitives de Colmez. Les méthodes de Zarhin et Colmez sont assez proches. Dans un premier temps, Zarhin montre qu'il y a une manière canonique d'associer une primitive f_ω à une 1-forme ω invariante sur un groupe algébrique commutatif sur K. Dans un deuxième temps, il utilise que toute 1-forme rationnelle fermée sur une variété lisse, vue sur son ouvert de définition, est l'image inverse d'une 1-forme invariante sur un groupe algébrique et il prend l'image inverse de la primitive définie sur ce groupe.

1.2. La construction de Coleman

Dans cette partie, on suppose pour simplifier $K = \mathbf{C}_p$.

1.2.1. *Quelques mots d'introduction sur la géométrie rigide.* On renvoie à [Ber], [Ga] ou [Ral] pour des introductions efficaces et concises à la géométrie rigide. Mais d'abord, pourquoi la géométrie rigide pour définir une intégration p-adique ? Essentiellement pour deux raisons. La première est que les fonctions rigides localement constantes sur un espace analytique rigide connexe sont automatiquement constantes. Ceci joue un rôle clef dans les démontrations ci-dessous. La deuxième est que les morphismes de Frobenius sur les schémas en caractéristique p se relèvent localement sur les *complétés p-adiques* des modèles lisses en caractéristique 0 (lorsque de tels modèles existent), mais pas sur les modèles eux-mêmes. On obtient ainsi des morphismes de Frobenius sur les fibres génériques au sens de Raynaud de ces complétés ([Ber],0.1.2), qui sont seulement des espaces rigides. Or, dans la théorie de Coleman, ces morphismes de Frobenius jouent un rôle analogue au théorème du carré dans l'intégrale de Colmez.

Rappelons qu'une algèbre de Tate (sur $\mathbf{C}_p$) est un quotient de $\mathbf{C}_p\{z_1,...,z_d\} = \{\sum_{\underline{\alpha}=(\alpha_1,...,\alpha_d)} a_{\underline{\alpha}} z_1^{\alpha_1}...z_d^{\alpha_d},\ a_{\underline{\alpha}} \to 0 \text{ si } |\underline{\alpha}| \to +\infty\}$ pour un d convenable. Si A est une algèbre de Tate, on note $\mathrm{Spm}(A) = \mathrm{Spec}(A)(\mathbf{C}_p)$ l'ensemble des idéaux maximaux de A : les $\mathrm{Spm}(A)$ jouent en géométrie rigide le rôle que jouent les schémas affines en géométrie algébrique. On veut les munir d'une "topologie" plus fine que la topologie de Zariski, mais pour laquelle les seules fonctions localement constantes sont les constantes (du moins si $\mathrm{Spec}(A)$ est connexe). Il n'est pas difficile de trouver des sous-ensembles de $\mathrm{Spm}(A)$ dont on a envie qu'ils soient des ouverts, par exemple $\mathrm{Spm}\big(A\{z\}/(zf-c)\big)$ où $f \in A$ et $c \in \mathbf{C}_p^*$, mais la topologie engendrée par ces ouverts redonne en général la topologie p-adique totalement discontinue, dont on ne veut pas. L'idée est alors de "sélectionner" certains ouverts et certains recouvrements, que l'on appelle admissibles, et remplacer la notion de topologie par la notion plus générale de topologie de Grothendieck. En munissant l'ensemble $\mathrm{Spm}(A)$ de cette topologie de Grothendieck et du faisceau structural rigide (dont les sections globales sur $\mathrm{Spm}(A)$ ne sont autres que A), on obtient l'espace analytique rigide affinoïde associé à $\mathrm{Spm}(A)$ (voir [Ber],0.1). Les espaces analytiques rigides généraux sont ensuite définis comme les ensembles munis d'une topologie de Grothendieck et d'un faisceau de $\mathbf{C}_p$-algèbres qui sont localement isomorphes à un espace rigide affinoïde. On définit des morphismes d'espaces rigides, des différentielles rigides (en particulier des 1-formes), etc... : voir ([Ber],0).

Exemple 1.2.1.1. — Si $A = \mathbf{C}_p\{z\}$, $\mathrm{Spm}(A) \simeq \{z \in \mathbf{C}_p \mid |z| \leq 1\}$, si $A = \mathbf{C}_p\{z_1, z_2\}/(z_1 z_2 - p)$, $\mathrm{Spm}(A) \simeq \{z \in \mathbf{C}_p \mid \frac{1}{p} \leq |z| \leq 1\}$.

Si X est un schéma sur $\mathrm{Spec}(\mathbf{C}_p)$, disons localement de type fini, on peut lui associer un espace analytique rigide X^{rig} dont l'ensemble sous-jacent est formé des points fermés $X(\mathbf{C}_p)$ de X. Tout ouvert de Zariski U dans X fournit par exemple un ouvert admissible $U(\mathbf{C}_p)$ pour la topologie de Grothendieck sur $X(\mathbf{C}_p)$. Mais d'autres ouverts jouent un rôle crucial : les tubes. Supposons pour simplifier X propre et soit $\mathcal{X}$ un modèle propre et plat

sur Spec($\mathcal{O}_{\mathbf{C}_p}$) de fibre spéciale Y sur Spec($\overline{\mathbf{F}}_p$). Le critère valuatif de propreté nous dit $X(\mathbf{C}_p) \simeq \mathcal{X}(\mathcal{O}_{\mathbf{C}_p})$, on a donc une flèche de spécialisation $sp : X(\mathbf{C}_p) \to Y(\overline{\mathbf{F}}_p)$. Soient Z un sous-schéma de Y, et $]Z[= sp^{-1}(Z(\overline{\mathbf{F}}_p))$: on peut montrer que $]Z[$ est un ouvert admissible de $X^{rig} = X(\mathbf{C}_p)$, qu'on appelle le tube de Z dans X^{rig} ([Ber],1.1.1).

Exemple 1.2.1.2. — Si Z est un ouvert de Zariski affine de Y, $]Z[$ est un ouvert affinoïde de X^{rig} ([Ber],0.2.2.1). Si $Z = \{y\}$ est un point fermé de Y et si de plus $\mathcal{X}$ est lisse, $]Z[$ est isomorphe à la boule ouverte $B(0,1_-)^d = \{(z_1,...,z_d) \in K^d \mid |z_i| < 1, 1 \leq i \leq d\}$ où d est la dimension de Y (même genre d'arguments que pour (1.1.1.1)).

Si X est un schéma sur $\mathbf{C}_p$, on note X^{rig} l'espace analytique rigide associé. Si X est un espace rigide sur $\mathbf{C}_p$, on note $\mathcal{O}_X$ le faisceau des fonctions rigides sur X.

1.2.2. *Intégration des formes de seconde espèce.* Soit X une variété algébrique sur Spec($\mathbf{C}_p$) admettant un modèle propre et lisse $\mathcal{X}$ sur Spec($\mathcal{O}_{\mathbf{C}_p}$) et notons Y sa fibre spéciale sur $\overline{\mathbf{F}}_p$. Comme Y est de type fini sur $\overline{\mathbf{F}}_p$, Y provient d'un modèle défini sur une extension finie (non unique) $\mathbf{F}_{p^r}$ de $\mathbf{F}_p$. On appelle morphisme de Frobenius sur Y tout morphisme de $\overline{\mathbf{F}}_p$-schémas $Y \to Y$ provenant par extension des scalaires pour un certain r de la puissance $r^{ième}$ du Frobenius absolu sur un tel modèle. Si ϕ et ϕ' sont deux morphismes de Frobenius sur Y, il est clair qu'on a $\phi^s = \phi'^t$ pour des entiers s, t convenables. Par ailleurs, il résulte de [BO] qu'on a un isomorphisme :

$$H^1_{dR}(X) \simeq \mathbf{C}_p \otimes H^1_{cris}(Y/W(\overline{\mathbf{F}}_p))$$

où $W(\overline{\mathbf{F}}_p)$ désigne les vecteurs de Witt à coefficients dans $\overline{\mathbf{F}}_p$ et H^1_{cris} le H^1 cristallin. Tout morphisme de Frobenius ϕ sur Y induit donc par fonctorialité un morphisme $\mathbf{C}_p$-linéaire $\phi^* : H^1_{dR}(X) \to H^1_{dR}(X)$ dont on note P_ϕ le polynôme caractéristique. C'est une conséquence du théorème de pureté de Deligne qu'aucune des racines de P_ϕ n'est une racine de l'unité (voir [KM]).

Avant d'énoncer (et de démontrer) le théorème principal de Coleman, disons quelques mots sur la méthode. On se rappelle que, dans le cas où X est une variété abélienne, la stratégie de Colmez était de remarquer que puisque $m^*_{\{1,2\}}\omega - m^*_{\{1\}}\omega - m^*_{\{2\}}\omega + m^*_{\emptyset}\omega = dg_\omega$ où ω est une 1-forme de seconde espèce quelconque sur X et g_ω une fonction *rationnelle* sur X^3 (cf. 1.1.2.3), il est naturel de chercher une primitive localement analytique f_ω telle que $m^*_{\{1,2\}}f_\omega - m^*_{\{1\}}f_\omega - m^*_{\{2\}}f_\omega + m^*_{\emptyset}f_\omega$ soit une fonction *rationnelle* sur X^3. Revenons à la situation précédente et supposons pour simplifier qu'il existe un morphisme $\phi : X^{rig} \to X^{rig}$ qui relève un morphisme de Frobenius sur Y (via la flèche de spécialisation) et induit ϕ^* sur $H^1_{dR}(X^{rig}) = H^1_{dR}(X)$. Pour toute 1-forme de seconde espèce ω sur X, vue sur X^{rig}, on a donc $P_\phi(\phi^*)(\omega) = dg_\omega$ où g_ω est une fonction *rigide* définie sur un ouvert de Zariski de X^{rig}. La stratégie de Coleman (antérieure) consiste alors à chercher une primitive localement analytique f_ω telle que $P_\phi(\phi^*)(f_\omega)$ soit une fonction *rigide*. Bien

sûr, en général, ϕ n'est défini que localement sur X^{rig} et il faut recoller...

Si U est un ouvert de Zariski dans Y, on note comme précédemment $]U[\subset X(\mathbf{C}_p)$ l'ensemble des points fermés de X qui se spécialisent sur Y en des points fermés de U. Un résultat de Coleman ([Cole3],1.1) nous dit que si U est affine, tout morphisme de Frobenius $\phi_U : U \to U$ se relève (de façon non unique) en un morphisme d'espaces rigides $\phi_{]U[} :]U[\to]U[$. Fixons ω une 1-forme rationnelle de seconde espèce sur X. On peut montrer qu'il existe un recouvrement affine fini Y_i de Y tel que $\omega|_{]Y_i[} = \omega_i + df_i|_{]Y_i[}$ où $\omega_i \in \Omega^1_{]Y_i[/\mathbf{C}_p}$ est une 1-forme rigide fermée et $f_i \in \mathbf{C}_p(X)$. Faisons le choix :
1) d'un tel recouvrement, 2) d'une écriture $\omega_i + df_i|_{]Y_i[}$ sur chaque $]Y_i[$, 3) d'un morphisme de Frobenius ϕ sur Y qui préserve les Y_i, 4) de relevés ϕ_i de ϕ sur les affinoïdes $]Y_i[$. Rappelons que $\mathcal{O}_{]Y_i[}$ désigne le faisceau structural rigide sur $]Y_i[$.

Théorème 1.2.2.1. — *Avec les notations ci-dessus, soit U_ω l'ouvert de Zariski de X complémentaire des pôles de ω ; il existe à addition près d'une constante une unique fonction f_ω localement analytique sur $U_\omega(\mathbf{C}_p)$ telle que :*
1) $df_\omega = \omega$
2) *pour tout i, $(f_\omega - f_i)|_{]Y_i[}$ se prolonge en une fonction localement analytique sur $]Y_i[$ et $P_\phi(\phi_i^*)((f_\omega - f_i)|_{]Y_i[}) \in \Gamma(]Y_i[, \mathcal{O}_{]Y_i[})$.*
De plus, f_ω est indépendante des choix 1) à 4) ci-dessus.

Preuve. — Notons pour alléger $\mathcal{O}(]Y_i[) = \Gamma(]Y_i[, \mathcal{O}_{]Y_i[})$. Le schéma de preuve est le suivant : dans une première étape, on montre que les conditions $dg_i = \omega_i$ et $P_\phi(\phi_i^*)(g_i) \in \mathcal{O}(]Y_i[)$ déterminent une unique fonction $g_i =$"$(f_\omega - f_i)|_{]Y_i[}$" localement analytique sur $]Y_i[$ modulo une constante ; dans une seconde étape, on montre que $g_i + f_i$ et $g_j + f_j$ coïncident sur $]Y_i[\cap]Y_j[$ modulo une constante et dans la dernière étape, on montre que l'on peut ajuster les constantes pour vraiment recoller les $g_i + f_i$.

Première étape : détermination de $(f_\omega - f_i)|_{]Y_i[}$.
La flèche de fonctorialité $H^1_{dR}(X) \to H^1_{dR}(]Y_i[)$ envoie $P_\phi(\phi^*)$(classe de ω) sur la classe de $P_\phi(\phi_i^*)(\omega_i)$. Cette classe est donc nulle par choix de P_ϕ, i.e., $P_\phi(\phi_i^*)(\omega_i) \in d\mathcal{O}(]Y_i[)$. Montrons qu'il existe une unique fonction g_i localement analytique sur $]Y_i[$ (à addition près d'une constante) telle que :

$$\begin{aligned} dg_i &= \omega_i \\ P_\phi(\phi_i^*)(g_i) &\in \mathcal{O}(]Y_i[). \end{aligned}$$

Quitte à multiplier P_ϕ par une constante, on a $P_\phi(T) = a_0 + a_1 T + ... + a_{n-1}T^{n-1} + T^n \in \mathbf{C}_p[T]$ et $P_\phi(\phi_i^*)(g_i) = \sum_k a_k(g_i \circ \phi_i^k)$. Si $\Omega_i = (\omega_i, \phi_i^*\omega_i, ..., (\phi_i^{n-1})^*\omega_i)$, on voit en posant $G_i = (g_i, g_i \circ \phi_i, ..., g_i \circ \phi_i^{n-1})$ qu'il suffit de montrer qu'il existe une unique fonction (à constante près) localement analytique $G_i :]Y_i[\to \mathbf{C}_p^n$ telle que $dG_i = \Omega_i$ et

$G_i \circ \phi_i - MG_i \in \mathcal{O}(]Y_i[)^n$ où :

$$M = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \\ -a_0 & -a_1 & \dots & \dots & -a_{n-1} \end{pmatrix}.$$

Unicité : S'il y en a deux, leur différence D_i vérifie $dD_i = 0$, donc est localement constante, et $D_i \circ \phi_i - MD_i$ aussi. Mais $D_i \circ \phi_i - MD_i \in \mathcal{O}(]Y_i[)^n$ et comme $]Y_i[$ est connexe (i.e. ne peut s'écrire $]Y_i[= U \cup U'$ où U et U' sont deux ouverts admissibles non vides d'intersection vide), $D_i \circ \phi_i - MD_i$ est constant de valeur C_i. Par ailleurs, comme 1 n'est pas racine de P_ϕ, $1 - M \in GL_n(\mathbf{C}_p)$ et un calcul donnent, pour $k \geq 1$:

$$D_i \circ \phi_i^k - M^k D_i = (1 - M^k)(1 - M)^{-1} C_i. \tag{1}$$

Soient $y_i \in Y_i(\overline{\mathbf{F}}_p)$ et $]y_i[= sp^{-1}(y_i) \subset]Y_i[$ le tube correspondant, isomorphe (en tant qu'espace rigide) à une boule ouverte (cf. 1.2.1.2). Comme y_i est défini sur une extension finie de $\mathbf{F}_p$, il existe $m \in \mathbf{N}$ tel que $\phi_i^m(]y_i[) \subset]y_i[$. De plus, pour tout $x \in]y_i[$, $\phi_i^{ml}(x)$ converge dans $]y_i[$ quand l tend vers $+\infty$ vers un ϵ_{y_i} indépendant de x ([Dw], lemme 3.0) et on a $\phi_i^m(\epsilon_{y_i}) = \epsilon_{y_i}$. En faisant $m = k$ dans (1) appliqué au point ϵ_{y_i}, on obtient :

$$D_i(\epsilon_{y_i}) = (1 - M)^{-1} C_i. \tag{2}$$

Par ailleurs, comme D_i est localement constante, $D_i(\phi_i^{ml_x}(x)) = D_i(\epsilon_{y_i})$ pour $l_x \gg 0$. En faisant $k = ml_x$ dans (1) appliqué au point x, on obtient en utilisant (2) :

$$M^{ml_x}(D_i(x) - (1 - M)^{-1} C_i) = 0. \tag{3}$$

Comme $\mathbf{C}_p$ est algébriquement clos, $\phi_i^m :]y_i[\to]y_i[$ est en fait surjectif et tout $x \in]y_i[$ est dans l'image de ϕ_i^{ml} pour tout l. En faisant $k = ml$ dans (1), un calcul donne $D_i(x) - (1 - M)^{-1} C_i \in M^{ml}(\mathbf{C}_p^n)$ pour tout l. Donc $D_i(x) - (1 - M)^{-1} C_i \in \mathrm{Ker}(M^{ml_x}) \cap \bigcap_{l \in \mathbf{N}} Im(M^{ml}) = \{0\}$ (algèbre linéaire élémentaire), ce qui montre que D_i est bien une fonction constante.

Existence : Puisque $P_\phi(\phi_i^*)(\omega_i) \in d\mathcal{O}(]Y_i[)$, on a $H_i \in \mathcal{O}(]Y_i[)^n$ tel que $\phi_i^* \Omega_i - M\Omega_i = dH_i$. Soient $y_i \in Y_i(\overline{\mathbf{F}}_p)$ et $\epsilon_{y_i} \in]y_i[$ comme précédemment, comme Ω_i est fermée et $]y_i[^n$ est une boule ouverte, il existe une unique fonction G_{y_i} analytique sur $]y_i[^n$ telle que $dG_{y_i} = \Omega_i|_{]y_i[}$ et $G_{y_i}(\epsilon_{y_i}) = (1 - M^m)^{-1} \sum_{l=0}^{m-1} M^l H_i(\phi_i^{m-l-1}(\epsilon_{y_i}))$ (i.e. on fixe cette valeur en ϵ_{y_i}). Soit G_i l'unique fonction localement analytique sur $]Y_i[$ telle que $G_i|_{]y_i[} = G_{y_i}$ pour $y_i \in Y_i(\overline{\mathbf{F}}_p)$

(les $]y_i[$ sont disjoints deux à deux). On a $dG_i = \Omega_i$ et un calcul montre que :

$$\begin{aligned}(G_i \circ \phi_i - MG_i)(\epsilon_{y_i}) &= H_i(\epsilon_{y_i}) \\ d(G_i \circ \phi_i - MG_i) &= dH_i.\end{aligned}$$

La première égalité ayant lieu pour tout $y_i \in Y_i(\overline{\mathbf{F}}_p)$, on en déduit $G_i \circ \phi_i - MG_i = H_i \in \mathcal{O}(]Y_i[)^n$.

Deuxième étape : Comparaison des fonctions sur $]Y_i[\cap]Y_j[$.
Montrons que pour tous i, j, $(f_i + g_i)|_{]Y_i[\cap]Y_j[} = (f_j + g_j)|_{]Y_j[\cap]Y_i[} + c_{ij}$ où $c_{ij} \in \mathbf{C}_p$. En remarquant que ϕ_i et ϕ_j laissent stable $]Y_i[\cap]Y_j[=]Y_i \times_Y Y_j[$, on voit qu'il suffit de montrer que 1) $f_i + g_i$ est indépendant (modulo cte) de l'écriture $\omega|_{]Y_i[} = \omega_i + df_i$ et 2) $f_i + g_i$ est indépendant (modulo cte) du relevé ϕ_i choisi. Pour 1), soit $\omega|_{]Y_i[} = \omega_i' + df_i'$ une autre écriture et g_i' la primitive de ω_i' construite comme dans la première étape. Alors $\omega_i - \omega_i' = d(g_i - g_i')$ et $P_\phi(\phi_i^*)(g_i - g_i') \in \mathcal{O}(]Y_i[)$. Mais on a aussi $\omega_i - \omega_i' = d(f_i' - f_i)$, donc $f_i' - f_i$ est défini sur tout $]Y_i[$ et $P_\phi(\phi_i^*)(f_i' - f_i) \in \mathcal{O}(]Y_i[)$, d'où $f_i' - f_i = g_i - g_i'$ modulo constante par l'unicité dans la première étape appliquée à la forme $\omega_i - \omega_i'$. Pour 2), Coleman montre ([Cole3],1.2a) que si ϕ_i et ϕ_i' sont deux relevés du même ϕ, alors $P_\phi(\phi_i^*)(g_i) - P_\phi(\phi_i'^*)(g_i) \in \mathcal{O}(]Y_i[)$ d'où $P_\phi(\phi_i'^*)(g_i) \in \mathcal{O}(]Y_i[)$ ce qui entraîne 2) par unicité.

Troisième étape : recollement.
Reste donc à recoller les $f_i + g_i$, i.e. à trouver des constantes $c_i \in \mathbf{C}_p$ telles que $c_i - c_j = c_{ij}$ pour tout i, j, et on définit alors f_ω comme l'unique fonction telle que $f_\omega|_{]Y_i[} = f_i + g_i + c_i$. Mais puisque le schéma Y est connexe, $\cap_i Y_i \neq \emptyset$, donc $\cap_i]Y_i[\neq \emptyset$ et il est alors élémentaire que l'on peut ajuster les constantes c_i.

Montrons enfin que f_ω ne dépend pas du choix du morphisme ϕ sur Y. Il suffit pour cela de montrer que remplacer ϕ par ϕ^s ne change pas f_ω. Comme f_ω ne dépend pas des autres choix, on garde le même recouvrement, la même écriture et on remplace juste ϕ_i par ϕ_i^s sur $]Y_i[$. On vérifie que $P_{\phi^s}(T^s) = \Pi_{\zeta^s=1} P_\phi(\zeta T)$. Donc $P_\phi(\phi_i^*)(g_i) \in \mathcal{O}(]Y_i[)$ entraîne $P_{\phi^s}((\phi_i^s)^*)(g_i) \in \mathcal{O}(]Y_i[)$ et le résultat découle encore de l'unicité de g_i dans la première étape. □

Comme dans la première partie, on pose, si P et Q sont deux points de $X(\mathbf{C}_p)$ où ω est définie : $\int_P^Q \omega = f_\omega(Q) - f_\omega(P)$.

Corollaire 1.2.2.2. — *Avec les notations précédentes :*
(i) $\int_P^Q (\lambda\omega + \lambda'\omega') = \lambda \int_P^Q \omega + \lambda' \int_P^Q \omega'$ $(\lambda, \lambda' \in \mathbf{C}_p)$
(ii) $\int_P^Q \omega = f(Q) - f(P)$ *si* $\omega = df$ *où* $f \in \mathbf{C}_p(X)$
(iii) *si* $g : X' \to X$ *est un morphisme de variétés propres sur* $\mathrm{Spec}(\mathbf{C}_p)$ *ayant bonne réduction,* $\int_P^Q g^*\omega = \int_{g(P)}^{g(Q)} \omega$.

Preuve. — Les énoncés (i) et (ii) résultent de l'unicité modulo cte de f_ω. L'énoncé (iii) est facile si g provient d'un morphisme $\mathcal{X} \to \mathcal{X}'$ entre des modèles propres et lisses. Sinon, soient $Alb(\mathcal{X})$ et $Alb(\mathcal{X}')$ les modèles de Néron des variétés d'Albanese $Alb(X)$ et $Alb(X')$ associées à X et X' ([BLR]). Des propriétés du modèle de Néron on déduit un diagramme commutatif :

$$\begin{array}{ccc} X' & \xrightarrow{g} & X \\ \downarrow & & \downarrow \\ \mathcal{X}' & & \mathcal{X} \\ \downarrow & & \downarrow \\ Alb(\mathcal{X}') & \longrightarrow & Alb(\mathcal{X}) \end{array}$$

d'où (iii) en remarquant que toute 1-forme de seconde espèce sur X provient, modulo une forme exacte, d'une 1-forme de seconde espèce sur $Alb(X)$ (car $H^1_{dR}(X) \simeq H^1_{dR}(Alb(X))$). □

Corollaire 1.2.2.3. — *L'intégrale $\int_P^Q \omega$ ne dépend pas du modèle choisi et coïncide avec l'intégrale définie par Colmez.*

Preuve. — Utiliser (ii) et (iii) dans (1.2.2.2), le fait que les variétés auxiliaires utilisées par Colmez pour définir son intégrale (i.e. $Alb(X)$ et $Alb(X)^3$) ont toutes bonne réduction lorsque X a bonne réduction et l'unicité en (1.1.2.5). □

Remarque 1.2.2.4. — Si σ est un automorphisme continu de $\mathbf{C}_p$, de l'unicité de f_ω résulte aussi le fait que $\sigma(\int_P^Q \omega) = \int_{\sigma(P)}^{\sigma(Q)} \sigma^*\omega$ où $\sigma : \mathbf{C}_p \otimes_\sigma X \to X$. En particulier, si X provient d'une variété définie sur une extension finie K de $\mathbf{Q}_p$, si $\sigma \in Gal(\overline{\mathbf{Q}}_p/K)$, $P, Q \in X(K)$ et ω est défini sur K, on a bien $\int_P^Q \omega \in K$.

2. APPLICATIONS

2.1. L'accouplement des périodes p-adiques

Dans cette partie, K est une extension finie de $\mathbf{Q}_p$.

Nous allons appliquer ce qui précède, en suivant Colmez, aux périodes p-adiques des variétés abéliennes. Voir [Il] et la remarque (2.1.3.4) pour l'historique de ce sujet, cas particulier des conjectures de Fontaine. C'est Coleman qui le premier a eu l'idée dans [Cole2] d'utiliser l'intégration p-adique pour redéfinir les périodes p-adiques des variétés abéliennes avec bonne réduction, ou plutôt leur image dans $\mathbf{C}_p$.

Comme les intégrales p-adiques ne sont pas multivaluées, on s'attendrait normalement à ce qu'il n'existe pas de périodes en p-adique ! Et effectivement, si on veut des périodes à valeurs dans $\mathbf{C}_p$ associées aux 1-formes de $H^0(X, \Omega^1_{X/K})$ (comme sur les complexes), on

les trouve *a posteriori* toutes nulles (cf. 2.1.2.2). Le miracle est qu'il existe néanmoins des périodes mais à un niveau "supérieur", c'est-à-dire à valeurs dans une $\overline{\mathbf{Q}}_p$-algèbre B_{dR}^+, introduite par Fontaine, qui se surjecte seulement sur $\mathbf{C}_p$.

2.1.1. *Quelques rappels sur* B_{dR}^+. L'anneau B_{dR}^+ a déjà été introduit dans ce séminaire ([Il]), mais nous aurons besoin des détails de sa construction. Les références pour ce qui suit sont [Fo1], [Fo2].

Soit R l'ensemble des suites $x = (x^{(0)}, x^{(1)}, ..., x^{(n)}, ...)$ d'éléments de $\mathcal{O}_{\mathbf{C}_p}$ telles que $(x^{(n+1)})^p = x^{(n)}$. On munit R des lois $\cdot$ et $+$ en posant $x \cdot y = (x^{(n)}y^{(n)})_{n\in\mathbf{N}}$ et $x + y = (s^{(n)})_{n\in\mathbf{N}}$ où :

$$s^{(n)} = \lim_{m\to+\infty} (x^{(n+m)} + y^{(n+m)})^{p^m}$$

les hypothèses faisant que cela converge bien dans $\mathcal{O}_{\mathbf{C}_p}$. On montre que ces lois font de R un anneau commutatif intègre d'élément unité $1 = (1, 1, ...)$. De plus :

$$p \cdot 1 = \lim_{m\to+\infty} (\underbrace{1 + ... + 1}_{p \text{ fois}})^{p^m} = 0$$

donc R est un anneau de caractéristique p et le Frobenius $x = (x^{(n)}) \mapsto x^p = ((x^{(n)})^p)$ sur R est clairement un isomorphisme, i.e., R est parfait. Il est de plus muni d'une action naturelle de $Gal(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ via l'action sur $\mathcal{O}_{\mathbf{C}_p}$ et d'une valuation $\mathrm{val}(x) = \mathrm{val}(x^{(0)})$ pour laquelle il est séparé et complet de corps résiduel $R/\{x|\mathrm{val}(x) > 0\} \simeq \overline{\mathbf{F}}_p$.

Comme R est parfait, il est tentant de considérer les vecteurs de Witt $W(R)$ à coefficients dans R ([Se2]) : tout élément de $W(R)$ s'écrit de façon unique $\sum_{n=0}^{+\infty} p^n[x_n]$ où $x_n \in R$ et $[x_n]$ est son représentant multiplicatif dans $W(R)$. On montre alors qu'on a une surjection d'anneaux :

$$\begin{array}{rccc} \theta & : & W(R) & \to & \mathcal{O}_{\mathbf{C}_p} \\ & & \sum_{n=0}^{+\infty} p^n[x_n] & \mapsto & \sum_{n=0}^{+\infty} p^n x_n^{(0)} \end{array}$$

dont le noyau est un idéal principal engendré par $p - [\underline{p}]$ où $\underline{p} = (p^{(n)}) \in R$ est un élément tel que $p^{(0)} = p$. On remarque que $W(R)$ est complet pour la topologie $(p, \mathrm{Ker}(\theta))$-adique $= (p, [\underline{p}])$-adique.

Définition 2.1.1.1. — *L'anneau* B_{dR}^+ *est le complété de* $W(R)[\frac{1}{p}]$ *par rapport à l'idéal* $\mathrm{Ker}(\theta) = (p - [\underline{p}])$.

La surjection θ se prolonge en une surjection encore notée $\theta : B_{dR}^+ \to \mathbf{C}_p$. L'anneau B_{dR}^+ est de valuation discrète complet d'idéal maximal $\mathrm{Ker}(\theta) = (p-[\underline{p}])B_{dR}^+$ et de corps résiduel $B_{dR}^+/\mathrm{Ker}(\theta) \simeq \mathbf{C}_p$, c'est-à-dire qu'il s'identifie *non canoniquement* à $\mathbf{C}_p[[T]]$. On peut

considérer plusieurs topologies sur B_{dR}^+. Nous munissons ici B_{dR}^+ de celle pour laquelle les $p^m W(R)+(\mathrm{Ker}(\theta))^k$ forment une base de voisinage de 0 quand m et k décrivent $\mathbf{N}$, c'est-à-dire que $(x_n)_n$ tend vers 0 si et seulement si, pour tout $M \in \mathbf{N}$, $x_n \in p^M W(R)+(\mathrm{Ker}(\theta))^M$ pour $n \gg 0$. Pour cette topologie, B_{dR}^+ est séparé et complet. Il existe une action naturelle et continue de $Gal(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ sur B_{dR}^+ via l'action sur R, qui commute à θ. Il n'existe pas de section $\mathbf{C}_p \hookrightarrow B_{dR}^+$ de θ qui soit compatible à l'action de Galois, mais on peut montrer que $\overline{\mathbf{Q}}_p$ s'identifie canoniquement à la clôture algébrique de $\mathbf{Q}_p$ dans B_{dR}^+ et que le diagramme :

$$\begin{array}{ccc} \overline{\mathbf{Q}}_p & \hookrightarrow & B_{dR}^+ \\ \| & & \downarrow \theta \\ \overline{\mathbf{Q}}_p & \hookrightarrow & \mathbf{C}_p \end{array}$$

est commutatif et compatible à $Gal(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$. En fait, si on munit $\overline{\mathbf{Q}}_p$ de la topologie *induite* par celle de B_{dR}^+ (qui n'est pas la topologie p-adique), Colmez a montré que B_{dR}^+ n'est autre que le complété de $\overline{\mathbf{Q}}_p$ pour cette topologie. Autrement dit, $\overline{\mathbf{Q}}_p$ est dense dans B_{dR}^+ (cf. appendice de [Fo2]).

2.1.2. *Accouplement des périodes p-adiques : énoncé du théorème.* Soit X une variété lisse sur K de dimension d. On note encore $\theta : X(B_{dR}^+) \to X(\mathbf{C}_p)$ l'application induite par $\theta : B_{dR}^+ \to \mathbf{C}_p$ et on remarque qu'elle est surjective (choisir une section $\mathbf{C}_p \hookrightarrow B_{dR}^+$ de θ). Par ailleurs, comme $\overline{\mathbf{Q}}_p \hookrightarrow B_{dR}^+$, $X(\overline{\mathbf{Q}}_p) \hookrightarrow X(B_{dR}^+)$.

Lemme 2.1.2.1. — *Soit f une fonction localement analytique sur $X(\overline{\mathbf{Q}}_p)$ qui se prolonge en une fonction localement analytique sur $X(\mathbf{Q}_p)$, alors f se prolonge aussi canoniquement à $X(B_{dR}^+)$.*

Preuve. — Soit $x \in X(B_{dR}^+)$, on peut trouver un ouvert de Zariski $U_x = \mathrm{Spec}(\overline{\mathbf{Q}}_p[x_1, ..., x_n, z_1, ..., z_d]/(f_1, ..., f_n))$ de $X \otimes_K \overline{\mathbf{Q}}_p$ tel que la projection sur $z_1, ..., z_d$ induit un isomorphisme entre un sous-ensemble V de $U_x(B_{dR}^+)$ contenant x et $\{(z_1, ..., z_d) \in (B_{dR}^+)^d \mid |\theta(z_i)| < \delta,\ 1 \leq i \leq d\}$ pour un δ suffisamment petit dans $\mathbf{R}^+ - \{0\}$. Comme $X(\overline{\mathbf{Q}}_p) \cap V \simeq \{(z_1, ..., z_d) \in \overline{\mathbf{Q}}_p^d \mid |z_i| < \delta,\ 1 \leq i \leq d\}$, et comme f se prolonge en une fonction analytique au voisinage de $\theta(x) \in X(\mathbf{C}_p)$, on peut supposer f analytique sur $X(\overline{\mathbf{Q}}_p) \cap V$ quitte à se restreindre à un ouvert de V contenant x. En remarquant que si $x \in B_{dR}^+$, $(x^n)_{n \in \mathbf{N}}$ converge dans B_{dR}^+ si et seulement si $(\theta(x)^n)_{n \in \mathbf{N}}$ converge dans $\mathbf{C}_p$, on voit que la série formelle en $z_1, ..., z_d$ définissant f sur $X(\overline{\mathbf{Q}}_p) \cap V$ converge dans B_{dR}^+ sur tout V. □

Si ω est une 1-forme rationnelle fermée sur X et f_ω une des primitives construites en (1.1.1.5), on note encore f_ω la prolongée à $X(B_{dR}^+)$ (à valeurs maintenant dans B_{dR}^+) et $\int_P^Q \omega = f_\omega(Q) - f_\omega(P)$ si $P, Q \in X(B_{dR}^+)$. Avant d'énoncer le théorème sur l'accouplement des périodes p-adiques des variétés abéliennes, donnons un exemple simple (mais sortant *stricto sensu* du cadre des variétés abéliennes).

Exemple 2.1.2.2. — Supposons $X = \mathrm{Spec}(\mathbf{C}[z, \frac{1}{z}]) = \mathbf{G}_m$. Soient γ un générateur de $H_1(X(\mathbf{C}), \mathbf{Z}) = H_1(\mathbf{C}^*, \mathbf{Z})$, $\omega = \frac{dz}{z} \in H^0(X, \Omega^1_{X/\mathbf{C}})$ et $\varepsilon_n = e^{2i\pi/p^n} \in X(\mathbf{C})$. On sait que :

$$\int_\gamma \omega = p^n \int_1^{\varepsilon_n} \omega = p^n \int_0^{\frac{2\pi}{p^n}} \frac{de^{i\theta}}{e^{i\theta}} = 2i\pi.$$

C'est la période classique. Supposons maintenant $X = \mathrm{Spec}(K[z, \frac{1}{z}])$ et soient $\gamma = (\varepsilon_n)_n$ un générateur de $T_p(X) = \varprojlim \mu_{p^n}(\overline{\mathbf{Q}}_p)$ et $\omega = \frac{dz}{z} \in H^0(X, \Omega^1_{X/K})$. On a $f_\omega = \log$ et, par définition :

$$p^n \int_1^{\varepsilon_n} \omega = p^n \log(\varepsilon_n) = \log(\varepsilon_n^{p^n}) = 0 \ !$$

Considérons maintenant $\tilde{\varepsilon}_n = [(\varepsilon_{m+n})_{m \in \mathbf{N}}] \in X(B^+_{dR})$ et soit $t = \log(\tilde{\varepsilon}_0) = -\sum_{n=1}^{+\infty} \frac{(1-\tilde{\varepsilon}_0)^n}{n} \in B^+_{dR} - \{0\}$. Bien sûr, $\tilde{\varepsilon}_n \neq \varepsilon_n$ mais $\theta(\tilde{\varepsilon}_n) = \varepsilon_n$. De plus :

$$p^n \int_1^{\tilde{\varepsilon}_n} \omega = p^n \log(\tilde{\varepsilon}_n) = \log(\tilde{\varepsilon}_0) = t.$$

Plus généralement, si $x \in B^+_{dR}$, soit $m_k(x)$ le plus grand entier de $\mathbf{Z}$ tel que $x \in p^{m_k(x)} W(R) + (\mathrm{Ker}(\theta))^{k+1}$ et disons qu'une suite $(x_n)_n$ de points de B^+_{dR} est bornée si la suite $(m_k(x_n))_n$ est bornée pour tout k. On peut voir que la suite $(\varepsilon_n)_n$ ci-dessus vue dans B^+_{dR} n'est pas une suite bornée, par contre la suite $(\tilde{\varepsilon}_n)_n$ est évidemment bornée (tous les $m_k(\tilde{\varepsilon}_n)$ sont nuls). Soit $(\tilde{\varepsilon}'_n)_n$ une autre suite bornée quelconque de B^+_{dR} vérifiant seulement $\theta(\tilde{\varepsilon}'_n) = \varepsilon_n$, alors on peut montrer que la suite $p^n \int_1^{\tilde{\varepsilon}'_n} \omega = p^n \log(\tilde{\varepsilon}'_n)$ *converge toujours vers* t dans B^+_{dR}. Cet élément t, introduit par Fontaine, peut être vu comme l'analogue p-adique de $2i\pi$.

Ce qui suit va consister à transposer l'exemple ci-dessus au cadre des variétés abéliennes. Rappelons qu'une 1-forme de seconde espèce sur X est par définition une 1-forme rationnelle fermée sur X dont le résidu est nul, i.e. qui ne contient pas de termes en $\frac{df}{f}$ dans une de ses écritures, et que $H^1_{dR}(X)$ s'identifie au quotient des 1-formes de seconde espèce sur X par les différentielles des fonctions rationnelles. On suppose maintenant que X est une variété abélienne. Soient ω une 1-forme de seconde espèce sur X et $\gamma = (\gamma_n)_n \in T_p(X) = \varprojlim X[p^n](\overline{\mathbf{Q}}_p)$ où $X[p^n](\overline{\mathbf{Q}}_p) = \{x \in X(\overline{\mathbf{Q}}_p) \mid [p^n]x = 0\}$. En vertu de l'exemple précédent, on a envie de considérer $\lim_{n \to +\infty} p^n \int_0^{\tilde{\gamma}_n} \omega$ pour des "bons" relevés $\tilde{\gamma}_n$ de γ_n dans $X(B^+_{dR})$ par θ.

Définition 2.1.2.3. — *Si Y est une variété lisse sur K, une suite $(\tilde{\gamma}_n)_n$ de points de $Y(B^+_{dR})$ est dite bornée dans $Y(B^+_{dR})$ si pour tout recouvrement fini $(U_i)_{i \in I}$ de Y par des ouverts de Zariski, il existe une décomposition $\mathbf{N} = \cup_{i \in I} I_i$ telle que $\{\tilde{\gamma}_n, n \in I_i\} \subset U_i(B^+_{dR})$ et telle que pour toute fonction $f_i \in \Gamma(U_i, \mathcal{O}_{U_i})$, la suite $\{f(\tilde{\gamma}_n), n \in I_i\}$ est bornée dans B^+_{dR}* (cf. 2.1.2.2).

Si U est un ouvert de Zariski de X, une suite bornée dans $U(B^+_{dR})$ est bornée dans $X(B^+_{dR})$ mais la réciproque est fausse (prendre une suite qui s'accumule autour du fermé

complémentaire de U dans X). On peut maintenant énoncer le théorème. On choisit une suite bornée $(\tilde{\gamma}_n)_n$ de relevés de $(\gamma_n)_n$ et on espère que la suite $p^n \int_0^{\tilde{\gamma}_n} \omega$ converge dans B_{dR}^+ vers une limite qui n'est pas trop souvent nulle... Mais il faut modifier 0 et $\tilde{\gamma}_n$ pour être sûr qu'ils ne s'approchent pas trop des pôles de ω dans $X(B_{dR}^+)$.

Théorème 2.1.2.4. — *Soient X une variété abélienne sur K, ω une 1-forme rationnelle de seconde espèce sur X, U_ω l'ouvert de Zariski de X complémentaire des pôles de ω et $\gamma = (\gamma_n)_n \in T_p(X)$.*

(i) *Il existe une suite $(\tilde{\gamma}_n)_n$ bornée dans $X(B_{dR}^+)$ et une suite $(\delta_n)_n$ bornée dans $U_\omega(B_{dR}^+)$ telles que $\theta(\tilde{\gamma}_n) = \gamma_n$ et la suite $(\delta_n \oplus \tilde{\gamma}_n)_n$ est bornée dans $U_\omega(B_{dR}^+)$.*

(ii) *Si $(\tilde{\gamma}_n)_n$ et $(\delta_n)_n$ sont de telles suites, la suite $p^n \int_{\delta_n}^{\delta_n \oplus \tilde{\gamma}_n} \omega$ converge dans B_{dR}^+ vers un élément $\int_\gamma \omega$ qui ne dépend que de γ et de la classe de ω dans $H_{dR}^1(X)$.*

(iii) *L'application de $H_{dR}^1(X) \times T_p(X)$ dans B_{dR}^+ qui à (ω, γ) associe $\int_\gamma \omega$ est bilinéaire, commute à l'action de Galois et envoie $H^0(X, \Omega^1_{X/K}) \times T_p(X)$ dans $(p - [\underline{p}])B_{dR}^+ = \mathrm{Ker}(\theta)$.*

(iv) *L'accouplement "période" ci-dessus est non dégénéré, i.e. si $(\omega_1, ..., \omega_{2d})$ fournit une base de $H_{dR}^1(X)$ et $(\gamma_1, ..., \gamma_{2d})$ une base de $T_p(X)$, la matrice $(\int_{\gamma_j} \omega_i)_{1 \le i,j \le 2d}$ appartient à $GL_{2d}(B_{dR})$ où $B_{dR} = Frac(B_{dR}^+)$.*

2.1.3. *Esquisse de preuve et relations de Riemann p-adiques.* Nous ne disons rien sur la preuve de (i) qui est purement technique (voir [Colm2],II.3.1).

Passons à (ii). Remarquons d'abord que si $(x_n)_n$ est une suite de B_{dR}^+ telle que $(px_n - x_{n-1})_n$ est bornée, alors $p^{n-1}(px_n - x_{n-1})$ tend vers 0 dans B_{dR}^+ quand $n \to +\infty$ et donc $p^n x_n$ converge dans B_{dR}^+. Oublions dans un premier temps les δ_n et supposons $[p]\tilde{\gamma}_n = \tilde{\gamma}_{n-1}$. Pour la convergence, il suffit de voir que la suite $y_n = p \int_0^{\tilde{\gamma}_n} \omega - \int_0^{\tilde{\gamma}_{n-1}} \omega$ est bornée dans B_{dR}^+. Mais $y_n = (pf_\omega - [p]^* f_\omega)(\tilde{\gamma}_n) + \mathrm{cte}$ et $pf_\omega - [p]^* f_\omega \in K(X)$ (car $[p]$ sur X induit la multiplication par p sur $H_{dR}^1(X)$). Si la suite $\tilde{\gamma}_n$ est bornée dans l'ouvert où $pf_\omega - [p]^* f_\omega$ est défini, alors par définition, $(y_n)_n$ est bornée dans B_{dR}^+. Si elle n'est pas bornée dans cet ouvert, il faut la modifier avec des δ'_n pour qu'elle le devienne. Dans le cas général, on décompose $y_n = p \int_{\delta_n}^{\delta_n \oplus \tilde{\gamma}_n} \omega - \int_{\delta_{n-1}}^{\delta_{n-1} \oplus \tilde{\gamma}_{n-1}} \omega$ sous la forme $y_n = y_{1,n} + y_{2,n} + y_{3,n} + y_{4,n}$ où :

$$
\begin{aligned}
y_{1,n} &= p \int_{\delta_n}^{\delta_n \oplus \tilde{\gamma}_n} \omega - p \int_{\delta'_n}^{\delta'_n \oplus \tilde{\gamma}_n} \omega &&= p(\Delta^{[2]} f_\omega)(\delta'_n, \delta_n \ominus \delta'_n, \tilde{\gamma}_n) \\
y_{2,n} &= p \int_{\delta'_n}^{\delta'_n \oplus \tilde{\gamma}_n} \omega - \int_{[p]\delta'_n}^{[p]\delta'_n \oplus [p]\tilde{\gamma}_n} \omega &&= (pf_\omega - [p]^* f_\omega)(\delta'_n \oplus \tilde{\gamma}_n) - (pf_\omega - [p]^* f_\omega)(\tilde{\gamma}_n) \\
y_{3,n} &= \int_{[p]\delta'_n}^{[p]\delta'_n \oplus [p]\tilde{\gamma}_n} \omega - \int_{\delta_{n-1}}^{\delta_{n-1} \oplus [p]\tilde{\gamma}_n} \omega &&= (\Delta^{[2]} f_\omega)(\delta_{n-1}, [p]\delta'_n \ominus \delta_{n-1}, [p]\tilde{\gamma}_n) \\
y_{4,n} &= \int_{\delta_{n-1}}^{\delta_{n-1} \oplus [p]\tilde{\gamma}_n} \omega - \int_{\delta_{n-1}}^{\delta_{n-1} \oplus \tilde{\gamma}_{n-1}} \omega &&= f_\omega(\delta_{n-1} \oplus [p]\tilde{\gamma}_n) - f_\omega(\delta_{n-1} \oplus \tilde{\gamma}_{n-1})
\end{aligned}
$$

et où δ'_n est une suite bornée dans $U_\omega(B_{dR}^+) \cap ([p]^* U_\omega)(B_{dR}^+)$ telle que $\delta'_n \oplus \tilde{\gamma}_n$ est bornée dans ce même ouvert (il en existe). On peut déduire que les suites $(y_{1,n})_n$, $(y_{2,n})_n$ et $(y_{3,n})_n$

sont bornées du fait que les fonctions $\Delta^{[2]}f_\omega$ et $pf_\omega - [p]^*f_\omega$ sont rationnelles. Quant à $(y_{4,n})_n$, on montre qu'elle est bornée en utilisant $\theta(\delta_{n-1} \oplus [p]\tilde{\gamma}_n) = \theta(\delta_{n-1} \oplus \tilde{\gamma}_{n-1})$. Donc $p^n \int_{\delta_n}^{\delta_n \oplus \tilde{\gamma}_n} \omega$ converge dans B_{dR}^+. Si $(\delta_n)_n$ est remplacé par $(\delta''_n)_n$ satisfaisant les propriétés (i) en (1.1.1.5) (sans changer $\tilde{\gamma}_n$), alors la suite $(\delta'''_n)_n$ définie par $\delta'''_n = \delta_n$ si n est pair et $\delta'''_n = \delta''_n$ si n est impair vérifie encore ces propriétés, d'où on déduit que la limite est indépendante de $(\delta_n)_n$. Si on choisit un autre relevé borné $(\tilde{\gamma}'_n)_n$, la suite $(\delta_n \oplus \tilde{\gamma}'_n)_n$ est encore bornée dans $U_\omega(B_{dR}^+)$ ([Colm2],II.3.1) et le même raisonnement en "mélangeant" les suites $(\tilde{\gamma}_n)_n$ et $(\tilde{\gamma}'_n)_n$ montre que la limite est indépendante du relevé borné de $(\gamma_n)_n$ choisi. Enfin, si $\omega = df_\omega$ avec $f_\omega \in K(X)$, la suite $\int_{\delta_n}^{\delta_n \oplus \tilde{\gamma}_n} \omega = f_\omega(\delta_n \oplus \tilde{\gamma}_n) - f_\omega(\delta_n)$ est par définition bornée ce qui entraîne que $p^n \int_{\delta_n}^{\delta_n \oplus \tilde{\gamma}_n} \omega$ converge vers 0 dans B_{dR}^+. Ceci achève (ii).

Passons au (iii). La linéarité par rapport à ω vient de la linéarité de l'intégrale. Si $((\gamma_n)_n, (\gamma'_n)_n) \in T_p(X) \times T_p(X)$, on prend $\tilde{\gamma}_n \oplus \tilde{\gamma}'_n$ comme relevé de $\gamma_n \oplus \gamma'_n$, qui forme encore une suite bornée dans $X(B_{dR}^+)$. On peut trouver $(\delta_n)_n$ bornée dans $U_\omega(B_{dR}^+)$ telle que les trois suites $(\delta_n \oplus \tilde{\gamma}_n)_n$, $(\delta_n \oplus \tilde{\gamma}'_n)_n$ et $(\delta_n \oplus \tilde{\gamma}_n \oplus \tilde{\gamma}'_n)_n$ y sont encore bornées. On écrit :

$$p^n \int_{\delta_n}^{\delta_n \oplus \tilde{\gamma}_n \oplus \tilde{\gamma}'_n} \omega = p^n \int_{\delta_n}^{\delta_n \oplus \tilde{\gamma}_n} \omega + p^n \int_{\delta_n \oplus \tilde{\gamma}_n}^{\delta_n \oplus \tilde{\gamma}_n \oplus \tilde{\gamma}'_n} \omega$$

et on remarque que la suite $(\delta_n \oplus \tilde{\gamma}_n)_n$ vérifie les propriétés (i) par rapport à $(\tilde{\gamma}'_n)_n$. Par (ii), cela entraîne que $p^n \int_{\delta_n \oplus \tilde{\gamma}_n}^{\delta_n \oplus \tilde{\gamma}_n \oplus \tilde{\gamma}'_n} \omega$ converge dans B_{dR}^+ vers $\int_{\gamma'} \omega$, d'où la linéarité par rapport à γ. Enfin, si $\omega \in H^0(X, \Omega^1_{X/K})$, la primitive f_ω en (1.1.1.5) nulle en 0 est un homomorphisme de groupes (cf. 1.1.2.8) donc $p^n \int_{\delta_n}^{\delta_n \oplus \tilde{\gamma}_n} \omega = p^n f_\omega(\tilde{\gamma}_n) = f_\omega([p]^n \tilde{\gamma}_n)$. Or $\theta(f_\omega([p]^n \tilde{\gamma}_n)) = f_\omega([p]^n \gamma_n) = f_\omega(0) = 0$, d'où $p^n \int_{\delta_n}^{\delta_n \oplus \tilde{\gamma}_n} \omega \in \mathrm{Ker}(\theta)$.

Le plus délicat est (iv), qui découle de l'analogue p-adique des relations de Riemann entre les périodes. Rappelons que si U est un ouvert de Zariski de X, on a un isomorphisme $\Gamma(U, \mathcal{O}_U) \otimes_K H^0(X, \Omega^1_{X/K}) \tilde{\to} H^0(U, \Omega^1_{U/K})$ (car X est abélienne). Soit $(\omega_1, ..., \omega_d)$ une base sur K de $H^0(X, \Omega^1_{X/K})$, en passant à $U^{an} = U(K)$, on voit que si G est une fonction localement analytique sur $U(K)$, on a $dG = \sum_{i=1}^d G_i \omega_i$ où les G_i sont localement analytiques sur $U(K)$.

Proposition 2.1.3.1. — ([Colm2],II.1.19) *Soit D un diviseur de X. Il existe une fonction G_D localement analytique sur $X(K) - D(K)$ et ayant une singularité logarithmique le long de D telle que $\Delta^* G_D = G_D(x \oplus y \oplus z) - G_D(x \oplus y) - G_D(x \oplus z) - G_D(y \oplus z) + G_D(x) + G_D(y) + G_D(z)$ soit le logarithme d'une fonction rationnelle sur X^3. De plus, si $dG_D = \sum_{i=1}^d G_{D,i}\omega_i$, $\omega_{D,i} = dG_{D,i}$ est une 1-forme de seconde espèce sur X pour tout i.*

À la base de cette proposition est le théorème du cube qui dit que l'image inverse $\Delta^* D$ de D sur X^3 par $\Delta : X^3 \to X$ est le diviseur d'une fonction rationnelle. D'autre part, rappelons qu'à D est associé l'accouplement de Weil $\varepsilon_D : T_p(X) \times T_p(X) \to T_p(\mathbf{G}_m)$ (cf.

([Mu1],IV.20) par exemple). Si $(\gamma,\gamma')\in T_p(X)\times T_p(X)$, $\log[\varepsilon_D(\gamma,\gamma')]\in \mathbf{Z}_p t$ (voir 2.1.2.2 pour t) et l'accouplement :

$$\log[\varepsilon_D]: T_p(X)\times T_p(X)\to \mathbf{Z}_p t$$

est bilinéaire et antisymétrique.

Théorème 2.1.3.2. — *(Relations de Riemann p-adiques) Soit* $(\gamma,\gamma')\in T_p(X)\times T_p(X)$, *alors* :

$$\log[\varepsilon_D(\gamma,\gamma')]=\sum_{i=1}^{d}\Big(\int_\gamma \omega_i\int_{\gamma'}\omega_{D,i}-\int_{\gamma'}\omega_i\int_\gamma\omega_{D,i}\Big).$$

Voir ([Colm2],II.3.5) pour la preuve. En conséquence, on voit que si $\gamma_1,\dots,\gamma_{2d}$ est une base de $T_p(X)$ et si on note $\omega_{d+i}=\omega_{D,i}$ $(i\in\{1,\dots,d\})$, on a l'égalité matricielle :

$$\big(\log[\varepsilon_D(\gamma_i,\gamma_j)]\big)_{1\le i,j\le 2d}={}^t\Big(\int_{\gamma_j}\omega_i\Big)_{1\le i,j\le 2d}\cdot\begin{pmatrix}0 & I_d\\ -I_d & 0\end{pmatrix}\cdot\Big(\int_{\gamma_j}\omega_i\Big)_{1\le i,j\le 2d}$$

où i est le numéro de la ligne et j celui de la colonne. Comme on peut toujours choisir D tel que l'accouplement de Weil soit non dégénéré, i.e. tel que :

$$det\Big(\log[\varepsilon_D(\gamma_i,\gamma_j)]\Big)_{1\le i,j\le 2d}\in \mathbf{Q}_p^{\times}t\subset B_{dR}^{\times}$$

cela montre que, pour un tel D, $(\omega_1,\dots,\omega_{2d})$ fournit une base de $H^1_{dR}(X)$ et $\Big(\int_{\gamma_j}\omega_i\Big)_{1\le i,j\le 2d}$ $\in GL_{2d}(B_{dR})$.

Remarque 2.1.3.3. — Les fonctions G_D méritent le nom de fonctions de Green p-adiques. Les 1-formes $\omega_{D,i}$ sont déjà définies dans [Cole2]. De (2.1.2.4), on peut déduire des isomorphismes $B_{dR}\otimes_K H^i_{dR}(X)\xrightarrow{\sim} B_{dR}\otimes_{\mathbf{Q}_p} H^i_{\acute{e}t}(X_{\overline{K}},\mathbf{Q}_p)$ pour tout i (où $H^i(X)=\bigwedge^i H^1(X)$) et les relations de Riemann p-adiques sont une autre façon de dire que pour $i=2$ l'isomorphisme envoie la classe du diviseur D dans $H^2_{dR}(X)$ sur sa classe dans $H^2_{\acute{e}t}(X_{\overline{K}},\mathbf{Q}_p)$.

Remarque 2.1.3.4. — Les premiers théorèmes de comparaison entre $H^1_{dR}(X)$ et $H^1_{\acute{e}t}(X_{\overline{K}},\mathbf{Q}_p)$ pour une variété abélienne X sur K sont dus à Tate dans le cas de bonne réduction ([Ta1]) et Raynaud dans le cas général ([Bog]), qui démontrent la célèbre décomposition de Hodge-Tate :

$$\mathbf{C}_p\otimes_{\mathbf{Q}_p} H^1_{\acute{e}t}(X_{\overline{K}},\mathbf{Q}_p)\simeq\Big(\mathbf{C}_p(-1)\otimes_K H^0(X,\Omega^1_{X/K})\Big)\oplus\Big(\mathbf{C}_p\otimes_K H^1(X,\mathcal{O}_X)\Big).$$

Cette décomposition a ensuite été redémontrée par Fontaine dans le cas général ([Fo3]), puis par Coleman dans le cas de bonne réduction ([Cole2]) via son intégration p-adique exposée au (1.2). L'énoncé avec B^+_{dR} comme en (2.1.2.4) est plus fort (la décomposition ci-dessus s'en déduit facilement) et a d'abord été obtenu par Fontaine et Messing (cf. [Wi]). C'est un cas particulier de conjectures en tout degré de cohomologie dues à Fontaine (appendice de [Fo1]) qui sont maintenant démontrées en toute généralité (travaux

de Fontaine-Messing, Faltings, Kato, Tsuji, etc...) : voir l'exposé d'Illusie [Il] dans ce séminaire et [Ts] pour plus de détails. Mais les preuves, où il n'est pas du tout question d'intégrer des formes différentielles, sont souvent longues et techniques. La preuve de Colmez pour le H^1, fidèle à la stratégie classique sur les complexes, n'en est que plus élégante.

2.2. **Les polylogarithmes p-adiques**

Dans cette partie, $K = \mathbf{C}_p$.

Dans le cas des courbes avec bonne réduction, Coleman a développé une variante de sa construction en (1.2) (les "basic wide open spaces" en anglais) qui lui permet d'intégrer toutes les formes différentielles rationnelles, mais aussi les formes rigides et même, par itération, des formes localement analytiques pas forcément algébriques ou rigides, ce que ne permet pas l'intégrale de Colmez. Nous présentons ci-dessous un cas particulier de cette construction que nous utilisons ensuite pour définir les polylogarithmes p-adiques.

2.2.1. *Itération de l'intégration sur les courbes.* Soient donc X une courbe propre et lisse sur $\mathrm{Spec}(\mathbf{C}_p)$ qui admet un modèle $\mathcal{X}$ propre et lisse sur $\mathrm{Spec}(\mathcal{O}_{\mathbf{C}_p})$ et Y la fibre spéciale sur $\mathrm{Spec}(\overline{\mathbf{F}}_p)$.

Soient $S \subset X(\mathbf{C}_p)$ un ensemble fini de points fermés, $U = X - S$ et $\overline{S}$ l'image de S dans $Y(\overline{\mathbf{F}}_p)$. Soient $U^{rig} = U(\mathbf{C}_p)$ (= espace analytique rigide associé à U) et $W = sp^{-1}(Y(\overline{\mathbf{F}}_p) - \overline{S}) \subset U^{rig}$ où sp est la flèche de spécialisation. *On suppose que l'application $S \to \overline{S}$ est injective.* L'ouvert U^{rig} est alors un cas particulier de ce que Coleman appelle un "basic wide open space" ([CS],2.1) et W est un "affinoïde sous-jacent" (cf. 1.2.1.2). On a dans ce cas :

$$U^{rig} - W = \cup_{s \in S} C_s$$

où C_s est une couronne ouverte autour de $s \in S$ isomorphe à $\{z \in \mathbf{C}_p \mid 0 < |z| < 1\}$. Via cet isomorphisme, on pose pour $r \in]0,1[$:

$$U_r = W \cup \Big(\bigcup_{s \in S} C_{s,r} \Big) \subset U$$

où $C_{s,r} \subset C_s$ correspond à $\{z \in \mathbf{C}_p \mid r < |z| < 1\} \subset \{z \in \mathbf{C}_p \mid 0 < |z| < 1\}$. Quand r tend vers 1, les U_r forment une famille décroissante de voisinages de W. Ils dépendent des isomorphismes choisis, mais si $(U'_{r'})_{r'}$ est une famille obtenue à partir d'autres choix de coordonnées locales sur les C_s, et si $r' \in]0,1[$, il existe r suffisamment proche de 1 tel que $U_r \subset U_{r'}$, et réciproquement. Cela suffit pour que les résultats ci-dessous soient indépendants de la famille $(U_r)_r$.

Fixons un morphisme de Frobenius sur Y (cf. 1.2.2) qui envoie $\overline{S}$ dans $\overline{S}$.

Proposition 2.2.1.1. — (i) *Pour r suffisamment proche de 1, il existe un morphisme rigide $\phi : U_r \to U^{rig}$ tel que $\phi(W) \subset W$, $\phi(C_s \cap U_r) \subset C_s$ ($s \in S$) et qui commute via la flèche de spécialisation avec le morphisme de Frobenius fixé sur $Y(\overline{\mathbf{F}}_p)$.*

(ii) *Si N est un entier fixé et $n \in \{0, ..., N\}$, il existe r suffisamment proche de 1 tel que $\phi^n = \phi \circ ... \circ \phi$ (n fois) est encore bien défini de U_r dans U^{rig}.*

Voir ([CS],2.2) pour la preuve. Si n est un entier fixé et si $1 - r$ est suffisamment petit, on peut donc définir par fonctorialité : $(\phi^*)^n = (\phi^n)^* : \Omega^1_{U^{rig}/\mathbf{C}_p} \to \Omega^1_{U_r/\mathbf{C}_p}$ où $\Omega^1_{U^{rig}/\mathbf{C}_p}$ désigne les 1-formes (rigides) sur U^{rig} (resp. avec U_r).

Proposition 2.2.1.2. — *Il existe un polynôme $P_1(T) \in \mathbf{C}_p[T]$ dont aucune des racines n'est une racine de l'unité tel que, pour tout $\omega \in \Omega^1_{U^{rig}/\mathbf{C}_p}$, tout ϕ comme en (2.2.1.1) et tout r suffisamment proche de 1, $P_1(\phi^*)(\omega) \in d\Gamma(U_r, \mathcal{O}_{U_r})$.*

Indication de preuve. — L'inclusion d'espaces rigides $U_r \hookrightarrow U^{rig}$ induit un isomorphisme $H^1_{dR}(U^{rig}) \xrightarrow{\sim} H^1_{dR}(U_r)$ où $H^1_{dR}(U^{rig}) = \Omega^1_{U^{rig}/\mathbf{C}_p}/d\Gamma(U^{rig}, \mathcal{O}_{U^{rig}})$ (resp. avec U_r). Le morphisme $\phi : U_r \to U^{rig}$ donne donc un morphisme $\phi^* : H^1_{dR}(U^{rig}) \to H^1_{dR}(U^{rig})$. Le résultat vient du fait que le polynôme caractéristique de ϕ^* s'identifie au numérateur de la fonction ζ de la courbe affine $Y - \overline{S}$, dont les racines sont des nombres de Weil, i.e. de valeur absolue complexe (pour tout plongement dans $\mathbf{C}$) une puissance non nulle de p. □

Pour tout $s \in S$, soit $\Gamma(C_s, \mathcal{O}_{C_s})_{log} = \Gamma(C_s, \mathcal{O}_{C_s})[\log(f), f \in \Gamma(C_s, \mathcal{O}_{C_s})^\times]$.

Théorème 2.2.1.3. — *Soient ϕ, P_1 et r comme ci-dessus et ω dans $\Omega^1_{U^{rig}/\mathbf{C}_p}$, il existe à addition près d'une constante une unique fonction f_ω localement analytique sur $U^{rig} = U(\mathbf{C}_p)$ telle que :*

1) $df_\omega = \omega$

2) $P_1(\phi^*)(f_\omega) \in \Gamma(U_r, \mathcal{O}_{U_r})$

3) *pour tout $s \in S$, $f_\omega|_{C_s} \in \Gamma(C_s, \mathcal{O}_{C_s})_{log}$.*

De plus f_ω est indépendant de r, ϕ, P_1 et du morphisme de Frobenius sur Y.

Preuve. — Soit $g \in \Gamma(U_r, \mathcal{O}_{U_r})$ tel que $P_1(\phi^*)(\omega) = dg$. En procédant comme dans la première étape de (1.2.2.1), mais avec l'affinoïde W au lieu de l'affinoïde $]Y_i[$, on a une unique (à cte près) fonction f_W localement analytique sur W telle que $df_W = \omega|_W$ et $P_1(\phi^*)(f_W) = g|_W \in \Gamma(W, \mathcal{O}_W)$. Sur C_s, il est facile de voir que tout 1-forme de $\Omega^1_{C_s/\mathbf{C}_p}$ s'intègre de façon unique (mod. cte) dans $\Gamma(C_s, \mathcal{O}_{C_s})_{log}$: soit f_s une telle primitive de $\omega|_{C_s}$. Soit f l'unique fonction localement analytique sur U^{rig} telle que $f|_W = f_W$ et $f|_{C_s} = f_s$ pour tout $s \in S$: il faut ajuster les constantes d'intégration pour avoir $P_1(\phi^*)(f) \in \Gamma(U_r, \mathcal{O}_{U_r})$. Soit $c_s = (P_1(\phi^*)(f_s) - g)|_{C_s \cap U_r} \in \Gamma(C_s \cap U_r, \mathcal{O}_{C_s \cap U_r})_{log}$ (on utilise ici $\phi(C_s \cap U_r) \subset C_s$), on a $dc_s = 0$ sur $C_s \cap U_r = C_{s,r}$ qui est connexe, donc c_s est constant. Soit f_ω tel que $f_\omega|_W = f_W$ et $f_\omega|_{C_s} = f_s - \frac{c_s}{P_1(1)}$: on vérifie que $P_1(\phi^*)(f_\omega) = g \in \Gamma(U_r, \mathcal{O}_{U_r})$, et que g détermine f_ω de façon unique. On renvoie à ([CS],2.3,2.4) pour l'indépendance par

rapport à P_1 et au relevé ϕ (même méthode pour ce dernier qu'en (1.2.2.1)). L'indépendance par rapport au morphisme de Frobenius sur Y se prouve comme dans (1.2.2.1). □

Remarquons que si ω est une forme rationnelle quelconque sur X (i.e. pas forcément de seconde espèce), $S \subset X(\mathbf{C}_p)$ l'ensemble des points où ω a un pôle, et si *de plus* l'application de spécialisation est injective sur S, (2.2.1.3) fournit une primitive de ω bien définie modulo constante. On peut montrer que cette primitive coïncide avec celle en (1.1.1.5), en particulier $f_\omega = \log(f)$ si $\omega = \frac{df}{f}$ avec f non nul sur U ([CS],2.5.1).

Nous allons itérer ce processus. Notons $\mathcal{O}_0(U^{rig}) = \Gamma(U^{rig}, \mathcal{O}_{U^{rig}})$ et $\mathcal{O}_0(U_r) = \Gamma(U_r, \mathcal{O}_{U_r})$ pour $0 < r < 1$. On définit les sous-espaces suivants des fonctions localement analytiques sur U^{rig} ou U_r :

$$\begin{aligned} \mathcal{O}_1(U^{rig}) &= \mathcal{O}_0(U^{rig}) + \sum_{\omega \in \Omega^1_{U^{rig}/\mathbf{C}_p}} \mathcal{O}_0(U^{rig}) f_\omega \\ \mathcal{O}_1(U_r) &= \mathcal{O}_0(U_r) + \sum_{\omega \in \Omega^1_{U^{rig}/\mathbf{C}_p}} \mathcal{O}_0(U_r) f_\omega|_{U_r} \end{aligned}$$

où f_ω est "la" primitive en (2.2.1.3) et r est suffisamment proche de 1.

Proposition 2.2.1.4. — (i) *Soit $f \in \mathcal{O}_1(U_r)$ (resp. $\mathcal{O}_1(U^{rig})$) tel que $df = 0$, alors f est constante sur U_r (resp. U^{rig}).*

(ii) *Il existe un polynôme $P_2(T) \in \mathbf{C}_p[T]$ dont aucune des racines n'est une racine de l'unité tel que, pour tout $\omega \in \mathcal{O}_1(U^{rig}) \otimes_{\mathcal{O}_0(U^{rig})} \Omega^1_{U^{rig}/\mathbf{C}_p}$, tout ϕ comme en (2.2.1.1) et tout r suffisamment proche de 1, $P_2(\phi^*)(\omega) \in d\mathcal{O}_1(U_r)$.*

Indication de preuve. — On renvoie à ([CS],2.4.4) pour (i). Pour (ii), soit $H^1_{dR}(\mathcal{O}_1(U^{rig}))$ $= (\mathcal{O}_1(U^{rig}) \otimes \Omega^1_{U^{rig}/\mathbf{C}_p})/d\mathcal{O}_1(U^{rig})$, l'application :

$$\begin{array}{ccc} \Omega^1_{U^{rig}/\mathbf{C}_p} \times \Omega^1_{U^{rig}/\mathbf{C}_p} & \to & H^1_{dR}(\mathcal{O}_1(U^{rig})) \\ (\omega, \omega') & \mapsto & f_\omega \omega' \mod d\mathcal{O}_1(U^{rig}) \end{array}$$

est bien définie et Coleman et de Shalit montrent ([CS],2.4.6) qu'elle induit un isomorphisme compatible à ϕ^* :

$$H^1_{dR}(U^{rig}) \otimes_{\mathbf{C}_p} H^1_{dR}(U^{rig}) \xrightarrow{\sim} H^1_{dR}(\mathcal{O}_1(U^{rig})).$$

Les valeurs propres de ϕ^* sur $H^1_{dR}(\mathcal{O}_1(U^{rig}))$ sont donc des produits de nombres de Weil (cf. preuve de (2.2.1.2)) : ce ne sont en particulier jamais des racines de l'unité et on peut prendre pour P_2 le polynôme caractéristique de ϕ^* sur $H^1_{dR}(\mathcal{O}_1(U^{rig}))$. □

Il est alors formel à partir de (2.2.1.4) de généraliser la preuve de (2.2.1.3) (et ce qu'elle utilise de (1.2.2.1)) au cas où $\omega \in \mathcal{O}_1(U^{rig}) \otimes \Omega^1_{U^{rig}/\mathbf{C}_p}$ en remplaçant $P_1(\phi^*)$ par $P_2(\phi^*)$ et $\Gamma(U_r, \mathcal{O}_{U_r}) = \mathcal{O}_0(U_r)$ par $\mathcal{O}_1(U_r)$. On définit de manière analogue $\mathcal{O}_2(U^{rig})$ comme le

$\mathcal{O}_1(U^{rig})$-module engendré par les constantes et les primitives qu'on vient d'obtenir des 1-formes de $\mathcal{O}_1(U^{rig}) \otimes \Omega^1_{U^{rig}/\mathbf{C}_p}$, et $\mathcal{O}_2(U_r)$ comme le $\mathcal{O}_1(U_r)$-module engendré par les constantes et les restrictions à U_r de ces primitives. La proposition (2.2.1.4) est encore valable avec $\mathcal{O}_2(U_r)$ ou $\mathcal{O}_2(U^{rig})$ et ainsi de suite... Par récurrence, on obtient finalement des suites croissantes canoniques d'espaces de fonctions localement analytiques : $\mathcal{O}_k(U^{rig}) \subset \mathcal{O}_{k+1}(U^{rig})$ et $\mathcal{O}_k(U_r) \subset \mathcal{O}_{k+1}(U_r)$ et des polynômes $P_k(T)$ (non uniques mais dont les racines ne sont pas des racines de l'unité) tels qu'on ait, si r est suffisamment proche de 1 :

Théorème 2.2.1.5. — *Soient $k \in \mathbf{N}$ et $\omega \in \mathcal{O}_k(U^{rig}) \otimes_{\mathcal{O}_0(U^{rig})} \Omega^1_{U^{rig}/\mathbf{C}_p}$, il existe à addition près d'une constante une unique fonction $f_\omega \in \mathcal{O}_{k+1}(U^{rig})$ telle que :*

1) $df_\omega = \omega$
2) $P_{k+1}(\phi^*)(f_\omega) \in \mathcal{O}_k(U_r)$
3) *pour tout $s \in S$, $f_\omega|_{C_s} \in \Gamma(C_s, \mathcal{O}_{C_s})_{log}$.*

Bien sûr, la récurrence ci-dessus peut se formaliser et se généraliser en dégageant les propriétés des $\mathcal{O}_k(U^{rig})$: c'est ce qui est fait dans [Cole1] et [CS] où sont introduits des "F-cristaux logarithmiques".

2.2.2. *Construction des polylogarithmes p-adiques.* Pour une introduction aux polylogarithmes complexes, on renvoie à l'exposé d'Oesterlé dans ce séminaire ([Oe]). Soit k un entier, rappelons que les polylogarithmes classiques $\ell_k(z)$ (plus exactement leur détermination principale) sont les fonctions holomorphes sur $\mathbf{C} - [1, +\infty[$ données sur la boule unité ouverte par la formule :

$$\ell_k(z) = \sum_{n=1}^{+\infty} \frac{z^n}{n^k}.$$

On a $\ell_0(z) = \frac{z}{z-1}$, $\ell_1(z) = -\log(1-z)$ et $d\ell_k(z) = \ell_{k-1}(z)\frac{dz}{z}$. Nous allons définir leurs analogues p-adiques, que nous noterons encore ℓ_k, vu qu'il ne sera (presque) plus question des polylogarithmes complexes dans la suite.

Avec les notations de la partie précédente, considérons :

$$\begin{aligned} X &= \mathbf{P}^1 \\ S &= \{1, \infty\} \subset \mathbf{P}^1(\mathbf{C}_p). \end{aligned}$$

On a :

$$\begin{aligned} U^{rig} &= X(\mathbf{C}_p) - S = \mathbf{C}_p - \{1\} \\ W &= \{z \in \mathbf{C}_p \mid |z| \leq 1,\ |z-1| \geq 1\} \\ C_1 &= \{z \in \mathbf{C}_p \mid 0 < |z-1| < 1\} \\ C_\infty &= \{z \in \mathbf{C}_p \mid |z| > 1\} \\ U_r &= \{z \in \mathbf{C}_p \mid |z| < \frac{1}{r},\ |z-1| > r\} \quad (0 < r < 1). \end{aligned}$$

Remarquons que $\mathcal{O}_0(U^{rig}) = \mathcal{O}_0(\mathbf{C}_p - \{1\})$ est l'ensemble des séries de Laurent $\sum_{n=-\infty}^{+\infty} a_n(z-1)^n$ qui convergent sur $\mathbf{C}_p - \{1\}$ et $\Omega^1_{U^{rig}/\mathbf{C}_p} = \mathcal{O}_0(U^{rig})dz$. On choisit comme morphisme de Frobenius $\phi(z) = z^p$ qui préserve bien S et envoie U_r dans U^{rig} et $U_r \cap C_s$ dans C_s ($s \in S$) lorsque $\frac{1}{p^{\frac{1}{p-1}}} < r < 1$ (car $\{\zeta \mid \zeta^p = 1\} \cap U_r = \emptyset$ pour de tels r). D'après le paragraphe précédent, on dispose des espaces de fonctions localement analytiques $\mathcal{O}_k(\mathbf{C}_p - \{1\})$ pour $k \in \mathbf{N}$.

Corollaire 2.2.2.1. — *Il existe une unique suite de fonctions $\ell_k(z) \in \mathcal{O}_k(\mathbf{C}_p - \{1\})$ telle que :*

$$\begin{aligned} \ell_0(z) &= \frac{z}{z-1} \\ \ell_k(0) &= 0 \\ d\ell_k(z) &= \frac{\ell_{k-1}(z)}{z}dz. \end{aligned}$$

De plus, $\ell_1(z) = -\log(1-z)$ et, si $|z| < 1$, $\ell_k(z) = \sum_{n=1}^{+\infty} \frac{z^n}{n^k}$.

Preuve. — L'existence et l'unicité sont conséquences de (2.2.1.5). Un examen des preuves de (1.2.2.1) et (2.2.1.3) montre que les fonctions de $\mathcal{O}_k(\mathbf{C}_p - \{1\})$ sont toutes analytiques sur la boule unité ouverte $\{z \in \mathbf{C}_p \mid |z| < 1\} \subset W$. Or, la fonction $\sum_{n=1}^{+\infty} \frac{z^n}{n^k}$ est aussi analytique sur cette boule et satisfait les propriétés de l'énoncé. On a donc $\ell_k(z) = \sum_{n=1}^{+\infty} \frac{z^n}{n^k}$ pour $|z| < 1$ par unicité (mod. cte) de la primitive analytique des 1-formes analytiques sur la boule ouverte. Vérifions que $P_1(T) = p - T$ satisfait $P_1(\phi^*)(\omega) \in d\mathcal{O}_0(U_r)$ pour tout $\omega \in \mathcal{O}_0(U^{rig})dz$ et tout $r \in]\frac{1}{p^{\frac{1}{p-1}}}, 1[$. On voit qu'il suffit de vérifier $P_1(\phi^*)(\frac{dz}{z-1}) \in d\mathcal{O}_0(U_r)$. Mais on a :

$$P_1(\phi^*)\Big(\frac{dz}{z-1}\Big) = p\Big(\frac{1}{z-1} - \frac{z^{p-1}}{z^p-1}\ Big)dz$$

et un calcul montre que cette 1-forme se prolonge au point ∞ (faire $z \leftrightarrow 1/z$), donc à $\mathbf{P}^1(\mathbf{C}_p) - \{z \in \mathbf{C}_p \mid |z-1| \leq r\}$ qui contient U_r et est isomorphe à une boule ouverte. Or,

sur une boule ouverte, toute 1-forme analytique fermée s'intègre, et donc en particulier $P_1(\phi^*)(\frac{dz}{z-1})$, i.e. $P_1(\phi^*)(\frac{dz}{z-1}) \in d\mathcal{O}_0(U_r)$. Pour montrer que $\ell_1(z) = -\log(1-z)$, il suffit d'avoir $P_1(\phi^*)(-\log(1-z)) \in \mathcal{O}_0(U_r)$. Mais :

$$P_1(\phi^*)(-\log(1-z)) = \log\left(1 - \left(1 - \frac{1-z^p}{(1-z)^p}\right)\right)$$

et pour $\frac{1}{p^{\frac{1}{p-1}}} < r < |z-1|$, on vérifie que $\left|1 - \frac{1-z^p}{(1-z)^p}\right| < 1$, donc :

$$P_1(\phi^*)(-\log(1-z)) = -\sum_{n\geq 1} \frac{\left(1 - \frac{1-z^p}{(1-z)^p}\right)^n}{n} \in \mathcal{O}_0(U_r). \qquad \square$$

Remarque 2.2.2.2. — On peut prendre pour $P_k(T)$ les polynômes $p^k - T$ pour tout k ([Cole1],5.2).

Voici une application des polylogarithmes p-adiques. Soient χ un caractère de Dirichlet non trivial et $L_p(s,\chi)$ la fonction L p-adique de Kubota-Leopoldt, obtenue par interpolation des valeurs aux entiers *négatifs* de la fonction $L(s,\chi)$ complexe de Dirichlet ([Iw],3). Il est naturel de se demander quelles sont alors les valeurs de $L_p(s,\chi)$ aux entiers *positifs*. La réponse est fournie par le théorème suivant, dû à Coleman ([Cole1]) :

Théorème 2.2.2.3. — *Supposons $p \neq 2$. Soient $\chi : (\mathbf{Z}/d\mathbf{Z})^\times \to \mathbf{C}_p^\times$ un caractère de Dirichlet de conducteur $d > 1$ et ω le caractère de Dirichlet "représentant de Teichmüller" : $(\mathbf{Z}/p\mathbf{Z})^\times \to \mathbf{C}_p^\times$. Pour tout $k \geq 1$, on a :*

$$L_p(k, \chi\omega^{1-k}) = \left(1 - \frac{\chi(p)}{p}\right)\frac{g(\chi,\zeta)}{d}\sum_{a=1}^{d-1}\overline{\chi}(a)\ell_k(\zeta^{-a})$$

où $\chi\omega^{1-k}$ est le caractère de Dirichlet produit des caractères χ et ω^{1-k}, ζ une racine primitive $d^{\text{ième}}$ quelconque de 1 et $g(\chi,\zeta)$ la somme de Gauss définie par $g(\chi,\zeta) = \sum_{a=1}^{d-1}\chi(a)\zeta^a$.

On a supposé $p \neq 2$ seulement pour simplifier l'exposition. Le cas $k = 1$, qui n'utilise que le logarithme p-adique, est dû à Leopoldt ([Iw],5). On remarquera l'analogie avec la formule classique :

$$L(k,\chi) = \frac{g(\chi,\zeta)}{d}\sum_{a=1}^{d-1}\overline{\chi}(a)\ell_k(\zeta^{-a})$$

(qui découle simplement de l'identité pour tout n : $d\chi(n) = \sum_{1\leq a,b\leq d-1}\chi(a)\overline{\chi}(b)\zeta^{a-bn}$) où ℓ_k désigne ici bien sûr le polylogarithme complexe.

Remarque 2.2.2.4. — Si l'on ne s'intéresse qu'aux polylogarithmes p-adiques, il existe une façon plus directe de les obtenir (voir [St]). Mais la méthode précédente (qui est celle originelle de Coleman) a l'avantage de s'étendre à des situations plus générales (voir *Régulateurs p-adiques* ci-dessous).

2.3. **Autres applications**

Nous passons en revue très rapidement certaines des autres applications de l'intégration p-adique.

Points de torsion sur les courbes
Dans [Cole3] et [Cole5], Coleman démontre, en utilisant l'intégration p-adique, la conjecture de Manin-Mumford, précédemment montrée par Raynaud ([Ra2]), qui dit que toute courbe de genre ≥ 2 contenue dans une variété abélienne sur un corps de caractéristique 0 n'a qu'un nombre fini de points de torsion. On se ramène au cas d'un corps de nombres et le point de départ de la preuve de Coleman est que si X est une courbe sur une extension finie K de $\mathbf{Q}_p$ avec bonne réduction et $\iota : X \to J$ un morphisme d'Albanese, $\iota(Q) - \iota(P)$ est de torsion dans J si et seulement si $\int_P^Q \iota^*\omega = 0$ sur X pour tout $\omega \in H^0(J, \Omega^1_{J/\mathbf{C}_p})$.

Régulateurs p-adiques
Dans [CS], Coleman et de Shalit généralisent l'intégration en (2.2.1) à des ouverts rigides un peu plus généraux que U^{rig} appelés "basic wide open spaces". Par une méthode tout à fait analogue à celle décrite en (2.2.1), cela leur permet d'intégrer sans restriction les logarithmes de toutes les fonctions rationnelles sur les courbes avec bonne réduction (et même sur certaines courbes à réduction semi-stable) et de définir ainsi des régulateurs p-adiques, généralisations des dilogarithmes p-adiques. Comme en (2.2.2.3), ils les utilisent pour calculer des valeurs spéciales de la fonction L p-adique associée à une courbe elliptique sur $\mathbf{Q}$ avec multiplication complexe et bonne réduction en p.

Loi de réciprocité et hauteurs p-adiques
Soit X une courbe algébrique propre et lisse sur $\mathbf{C}_p$. Rappelons qu'une forme de troisième espèce sur X est une forme rationnelle ayant au plus des pôles simples et tels que les résidus en ces pôles soient des entiers. Si ω est une telle forme, le diviseur des résidus $Res(\omega) = \sum_P n_P P$ où P parcourt $X(\mathbf{C}_p)$ et n_P est le résidu en P est de degré 0 (i.e. $\sum n_P = 0$). Pour toute forme rationnelle ω' dont les pôles évitent $Res(\omega)$, $\int_{Res(\omega)} \omega' = \sum_P n_P f_{\omega'}(P)$ ne dépend donc pas de la primitive $f_{\omega'}$ choisie en (1.1.1.5). Si η et η' sont deux formes de seconde espèce sur X, notons $\eta \cup \eta'$ la somme des résidus de $f_\eta \eta'$ où f_η est une primitive en (1.1.1.5) et où on calcule le résidu en un point à partir des développements analytiques de f_η et η' dans un voisinage p-adique de ce point (c'est indépendant des choix car η' a tous ses résidus nuls).
Soit J la jacobienne de X et $\tilde{J}$ l'extension universelle de J par un groupe additif ([Me]). Les points de $\tilde{J}(\mathbf{C}_p)$ représentent les formes de troisième espèce sur X modulo les différentielles logarithmiques des fonctions rationnelles et l'algèbre de Lie de $\tilde{J}$ est isomorphe à $H^1_{dR}(X)$. On a en particulier une application logarithme $Log_{\tilde{J}} : \tilde{J}(\mathbf{C}_p) \to H^1_{dR}(X)$.

Théorème 2.3.1. — *(Loi de réciprocité) Si ω et ω' sont deux formes différentielles de troisième espèce dont les diviseurs des résidus sont étrangers, alors :*

$$\int_{Res(\omega)} \omega' - \int_{Res(\omega')} \omega = Log_{\bar{J}}(\omega) \cup Log_{\bar{J}}(\omega').$$

Ce théorème a d'abord été démontré par Coleman dans le cas où X a bonne réduction ([Cole4]), puis par Colmez dans le cas général ([Colm2]).
Soit maintenant X une courbe propre et lisse sur un corps de nombres et J sa jacobienne. À partir des fonctions de Green p-adiques (cf. 2.1.3.1) Colmez définit un accouplement de hauteur bilinéaire et symétrique : $J(\overline{\mathbf{Q}}) \times J(\overline{\mathbf{Q}}) \to \Pi_{p \leq \infty} \mathbf{Q}_p$. La symétrie découle de (2.3.1). Projeté sur $\mathbf{R}$, cet accouplement coïncide avec l'accouplement de Néron-Tate. Projeté sur $\mathbf{Q}_p$ aux places p où X a bonne réduction, il avait déjà été défini par Coleman et Gross ([CG]). À noter que le fait de pouvoir considérer $\log(p)$ comme une *variable* est important dans la définition de l'accouplement ci-dessus.

Cohomologie de de Rham et réduction semi-stable
Si X est une courbe propre et lisse sur une extension finie K de $\mathbf{Q}_p$ ayant réduction semi-stable (au sens où elle admet un modèle $\mathcal{X}$ propre, plat et régulier sur les entiers de K dont la fibre spéciale est un diviseur à croisements normaux dans $\mathcal{X}$), et si de plus les composantes irréductibles de la fibre spéciale sont lisses, leurs tubes dans X^{rig} (cf. 1.2.1) sont des "basic wide open spaces" sur lesquels on peut intégrer les formes différentielles. En utilisant l'intégration sur ces tubes, Coleman et Iovita donnent dans [CI] une construction purement géométrique de la K_0-structure de Fontaine-Jannsen D définie par Hyodo-Kato ([HK]) sur $H^1_{dR}(X)$ (i.e. $H^1_{dR}(X) = K \otimes_{K_0} D$) où K_0 est l'extension maximale de $\mathbf{Q}_p$ non ramifiée dans K, ainsi que des opérateurs ϕ et N définis sur D (Frobenius et monodromie). Voir aussi les résultats de Le Stum à ce sujet ([Le]).

RÉFÉRENCES

[Ber] Berthelot P. – *Cohomologie rigide et cohomologie rigide à supports propres, première partie*, prépublication de l'Institut de Recherche Mathématique de Rennes, 1996.

[Bes] Besser A. – *A generalization of Coleman's p-adic integration theory*, prépublication du Mathematische Institut der Universität Münster, 1999.

[BO] Berthelot P. – Ogus A., *F-isocristals and de Rham cohomology I*, Inv. Math. 72, 1983, 159-199.

[Bog] Bogomolov F. – *Sur l'algébricité des représentations l-adiques*, C.R.A.S. 290, 1980, 701-703.

[BLR] Bosch S. – Lütkebohmert W., Raynaud M., *Néron models*, Ergebnisse der Mathematik und ihrer Grenzgebiete 21, Springer-Verlag, 1990.

[Bou] Bourbaki N. – *Groupes et algèbres de Lie, Chapitres* 2 *et* 3, Hermann, 1972.

[Colm1] Colmez P. – *Périodes p-adiques des variétés abéliennes*, Math. Ann. 292, 1992, 629-644.

[Colm2] Colmez P. – *Intégration sur les variétés p-adiques*, Astérisque 248, Soc. Math. de France, 1998.

[Cole1] Coleman R. – *Dilogarithms, regulators and p-adic L-functions*, Inv. Math. 69, 1982, 171-208.

[Cole2] Coleman R. – *Hodge-Tate periods and p-adic abelian integrals*, Inv. Math. 78, 1984, 351-379.

[Cole3] Coleman R. – *Torsion points on curves and p-adic abelian integrals*, Annals of Math. 121, 1985, 111-168.

[Cole4] Coleman R. – *Reciprocity laws on curves*, Comp. Math. 72, 1989, 205-235.

[Cole5] Coleman R. – *Ramified torsion points*, Duke Math. J. 54, 1987, 615-640.

[CG] Coleman R., Gross B. – *p-adic heights on curves*, Adv. Math. 17, 1989, 73-81.

[CI] Coleman R., Iovita A. – *The Frobenius and Monodromy operators for Curves and Abelian Varieties*, Duke Math. J. 97, 1999, 171-215.

[CS] Coleman R., de Shalit E. – *p-adic regulators on curves and special values of p-adic L-functions*, Inv. Math. 93, 1988, 239-266.

[Dw] Dwork B. – *Normalized period matrices I*, Annals of Math. 94, 1971, 337-388.

[Fo1] Fontaine J.-M. – *Sur certains types de représentations p-adiques du groupe de Galois d'un corps local; construction d'un anneau de Barsotti-Tate*, Annals of maths 115, 1982, 529-577.

[Fo2] Fontaine J.-M. – *Le corps des périodes p-adiques*, Astérisque 223, Soc. Math. de France, 1994, 59-111.

[Fo3] Fontaine J.-M. – *Formes différentielles et modules de Tate des variétés abéliennes sur les corps locaux*, Inv. Math. 65, 1982, 379-409.

[Ga] Garuti M. – *Géométrie rigide et géométrie formelle*, à paraître aux Proceedings de la conférence de Luminy sur le groupe fondamental des courbes en caractéristique $p > 0$, décembre 1998.

[Il] Illusie L. – *Cohomologie de de Rham et cohomologie étale p-adique (d'après G. Faltings, J.-M. Fontaine et al.)*, Sém. Bourbaki, Exposé n° 726, Astérisque 189-190, 1990, 325-374.

[Iw] Iwasawa K. – *Lectures on p-adic L-functions*, Annals of Math. Studies 74, Princeton University Press, 1972.

[KM] Katz N., Messing W. – *Some consequences of the Riemann hypothesis for varieties over finite fields*, Inv. Math. 23, 1974, 73-77.

[La] Lang S. – *Abelian varieties*, Interscience, New-York, 1959.

[Le] Le Stum B. – *La structure de Hyodo-Kato pour les courbes*, Rend. Sem. Mat. Univ. Padova 94, 1995, 279-301.

[Me] Messing W. – *The universal extension of an abelian variety by a vector group*, Symp. Math. XI (Rome 1972), Academic Press, 1973, 359-372.

[Mu1] Mumford D. – *Abelian varieties*, Oxford University Press, 1970.

[Mu2] Mumford D. – *Analytic construction of degenerating curves over complete local fields*, Compos. Math. 24, 1972, 129-174.

[Oe] Oesterlé J. – *Polylogarithmes*, Sém. Bourbaki, Exposé n° 762, Astérisque 216, 1993, 49-67.

[Ra1] Raynaud M. – *Géométrie analytique rigide, d'après Tate, Kiehl,...*, Mémoire de la S.M.F. 39-40, 1974, 319-327.

[Ra2] Raynaud M. – *Courbes sur une variété abélienne et points de torsion*, Inv. Math. 71, 1983, 207-233.

[Ra3] Raynaud M. – *Variétés abéliennes et géométrie rigide*, Proceedings du Congrès International de Nice, 1970, 473-477.

[Se1] Serre J.-P. – *Morphismes universels et variétés d'Albanese*, Séminaire Chevalley 1958/1959, exposé 10.

[Se2] Serre J.-P. – *Corps locaux*, $3^{ième}$ édition, Hermann, 1968.

[St] Stevens G. – *p-adic polylogarithms*, notes d'une conférence à Schloss Ringberg am Tegernsee, 4 septembre 1998.

[Ta1] Tate J. – *p-divisible groups*, Proc. of a conf. on local fields, Driebergen, 1967, 158-183.

[Ta2] Tate J. – *Rigid analytic spaces*, Inv. Math. 12, 1971, 257-289.

[Ts] Tsuji T. – *p-adic-étale cohomology and crystalline cohomology in the semi-stable reduction case*, à paraître à Inv. Math.

[Wi] Wintenberger J.-P. – *Théorème de comparaison p-adique pour les schémas abéliens I : construction de l'accouplement de périodes*, Astérisque 223, Soc. Math. de France, 1994, 349-397.

[Za1] Zarhin Y. – *Local heights and abelian varieties*, Groupe d'étude sur les problèmes diophantiens 1988/1989, Publ. math. n° 90 de l'université Pierre et Marie Curie.

[Za2] Zarhin Y. – *p-adic abelian integrals and commutative Lie groups*, Algebraic geometry 4, J. Math. Sci. 81, 1996, 2744-2750.

Christophe BREUIL

Mathématiques, Bât. 425
U.R.A. 752 du C.N.R.S.
Université Paris-Sud
Bâtiment 425
F-91405 ORSAY cedex
E-mail : breuil@math.u-psud.fr

Astérisque

ROBERT BRYANT

Recent advances in the theory of holonomy

Astérisque, tome 266 (2000), Séminaire Bourbaki,
exp. n° 861, p. 351-374

<http://www.numdam.org/item?id=SB_1998-1999__41__351_0>

© Société mathématique de France, 2000, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (http://smf4.emath.fr/Publications/Asterisque/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Séminaire BOURBAKI Juin 1999
51ème année, 1998-99, n° 861

RECENT ADVANCES IN THE THEORY OF HOLONOMY

by **Robert BRYANT**

1. INTRODUCTION

1.0. OUTLINE. This report is organized as follows:

1. Introduction
2. Riemannian Holonomy
3. Torsion-free non-metric connections: the irreducible case

In a short lecture of this nature, it is impossible to describe the history of the subject in any depth, but the reader can find more information on the Riemannian (and pseudo-Riemannian) case by consulting [Bes], [Sa], and the forthcoming, much-anticipated [Jo3], especially for the exceptional cases. For the non-metric case, aside from the survey [Br3], the expository papers [MS2] and [Schw] provide a valuable account of both the representation theoretic and twistor theoretic approaches to the study of holonomy.

1.1. HISTORICAL REMARKS. According to the Oxford English Dictionary, it was Heinrich Hertz in 1899 who introduced the words *holonomic* and *nonholonomic* to describe a property of velocity constraints in mechanical systems.

Velocity constraints are *holonomic* if they force a curve in state space to stay in a proper subspace. As an example, the condition $\mathbf{p}\cdot d\mathbf{p} = 0$ for a vector particle $\mathbf{p} \in \mathbb{R}^n$ forces $\mathbf{p}$ to have constant length, while the constraint $\mathbf{p}\wedge d\mathbf{p} = 0$ forces $\mathbf{p}$ to move on a line.

Nonholonomic constraints, on the other hand, imply no such 'finite' constraints. A classical example is that of a ball rolling on a table without slipping or twisting. The state space is $B = \mathrm{SO}(3)\times\mathbb{R}^2$, where the $\mathrm{SO}(3)$ records the orientation of the ball and the $\mathbb{R}^2$ records its contact point on the plane. The rolling constraint is expressed

as the set of differential equations

$$\alpha \coloneqq a^{-1}\,da + a^{-1}\begin{pmatrix} 0 & 0 & -dx \\ 0 & 0 & -dy \\ dx & dy & 0 \end{pmatrix} a = 0 \tag{0.1}$$

for a curve $\big(a(t); x(t), y(t)\big)$ in B. The curves in B satisfying this constraint are those tangent to the 2-plane field $D = \ker\alpha$ that is transverse to the fibers of the projection $\mathrm{SO}(3)\times\mathbb{R}^2 \to \mathbb{R}^2$. It is not difficult to show that any two points of B can be joined by a curve tangent to D.

Such constraints and their geometry have long been of considerable interest in the calculus of variations and control theory. For recent results, see the foundational work on Carnot-Carathéodory geometries by Gromov [Gr].

1.2. THE HOLONOMY GROUP. Élie Cartan [Ca2] introduced the *holonomy group* in the context of differential geometry. It measures the failure of the parallel translation associated to a connection to be holonomic.

The data are a connected manifold M, a Lie group H with Lie algebra $\mathfrak{h}$, a principal right H-bundle $\pi : B \to M$, and a connection θ on B, i.e., θ is an $\mathfrak{h}$-valued 1-form on B that pulls back to each π-fiber to be the canonical left invariant 1-form on H and that satisfies the equivariance relation $R_h^*(\theta) = \mathrm{Ad}(h^{-1})(\theta)$. (The nonholonomic example described above with $(M, H, B, \theta) = \big(\mathbb{R}^2,\ \mathrm{SO}(3),\ \mathrm{SO}(3)\times\mathbb{R}^2,\ \alpha\big)$ is an example.)

A piecewise C^1 curve $\gamma : [0,1] \to B$ is said to be *θ-horizontal* or *θ-parallel* if $\gamma^*(\theta) = 0$. The *θ-holonomy* $B_u^\theta \subset B$ of $u \in B$ is defined to be the set of values of $\gamma(1)$ as $\gamma : [0,1] \to B$ ranges over the θ-horizontal curves with $\gamma(0) = u$. The $\mathrm{Ad}(H)$-equivariance of θ (coupled with the connectedness of M) implies that there is a subgroup $H_u^\theta \subset H$ so that B_u^θ is an H_u^θ-subbundle of B and that $H_{u\cdot h}^\theta = h^{-1}\,H_u^\theta\,h$ for $h \in H$. Consequently, the conjugacy class of $H_u^\theta \subset H$ depends only on θ. The group H_u^θ (or, more informally, its conjugacy class in H) is called the *holonomy* of θ.

It is a theorem of Borel and Lichnerowicz [BL] that H_u^θ is a Lie subgroup of H. By a theorem of Ambrose and Singer [AS], the Lie algebra $\mathfrak{h}_u^\theta$ of H_u^θ is spanned by the set

$$\{\ \Theta(x,y)\ \big|\ \ x, y \in T_vB\,,\ v \in B_u^\theta\ \}$$

where $\Theta = d\theta + \frac{1}{2}[\theta,\theta]$ is the *curvature* of θ. The identity component $(H_u^\theta)^0 \subset H_u^\theta$ is known as the *restricted holonomy* of θ. There is a well-defined surjective homomorphism $\rho^\theta : \pi_1\big(M, \pi(u)\big) \to H_u^\theta/(H_u^\theta)^0$ that satisfies $\rho^\theta\big([\pi\circ\gamma]\big) = \gamma(1)(H_u^\theta)^0$ for every θ-horizontal curve γ with $\gamma(0) = u$ and $\gamma(1) \in u\cdot H$.

Using these results, it can be shown [KNo] that, when $\dim M > 1$, a Lie subgroup $G \subset H$ can be the holonomy group of a connection on B if and only if B admits a structural reduction to a G-bundle. Thus, the set of possible holonomies of connections on B is determined topologically.

1.3. H-STRUCTURES AND TORSION. A common source of connections in geometry is that of H-structures on manifolds. Suppose $\dim M = n$ and let $\mathfrak{m}$ be a reference vector space of dimension n. An ($\mathfrak{m}$-valued) *coframe* at $x \in M$ is a linear isomorphism $u : T_xM \to \mathfrak{m}$. The set $F(M, \mathfrak{m})$ of $\mathfrak{m}$-valued coframes at the points of M is naturally a principal right $\mathrm{GL}(\mathfrak{m})$-bundle over M, with basepoint projection $\pi : F(M, \mathfrak{m}) \to M$. There is a tautological $\mathfrak{m}$-valued 1-form ω on $F(M, \mathfrak{m})$ defined by the formula $\omega(v) = u\big(\pi'(v)\big)$ for all $v \in T_uF(M, \mathfrak{m})$.

Let $H \subset \mathrm{GL}(\mathfrak{m})$ be a subgroup. An *H-structure* on M is an H-subbundle B such that $B \subset F(M, \mathfrak{m})$. When H is a closed subgroup of $\mathrm{GL}(\mathfrak{m})$ (the only case I will consider today), the set of H-structures on M is the set of sections of the bundle $F(M, \mathfrak{m})/H \to M$. The problem of determining whether there exists an H-structure on M is a purely topological one.

Most of the familiar geometric structures on M can be described as H-structures. For example, when $H = \mathrm{O}(Q) \subset \mathrm{GL}(\mathfrak{m})$ is the group of linear transformations preserving a quadratic form Q of type (p, q) on $\mathfrak{m}$, a choice of H-structure on M is equivalent to a choice of pseudo-Riemannian metric of type (p, q) on M. When $H = \mathrm{Sp}(S) \subset \mathrm{GL}(\mathfrak{m})$ is the group of linear transformations preserving a nondegenerate skewsymmetric bilinear form S on $\mathfrak{m}$, a choice of H-structure on M is equivalent to a choice of a nondegenerate 2-form σ on M, i.e., an almost symplectic structure.

If $\pi : B \to M$ is an H-structure on M, the tautological form ω pulled back to B will also be denoted ω when there is no chance of confusion. If θ is a connection on B, then the *first structure equation of Cartan* says that there exists an H-equivariant function $T : B \to \mathrm{Hom}\big(\Lambda^2(\mathfrak{m}), \mathfrak{m}\big) \simeq \mathfrak{m} \otimes \Lambda^2(\mathfrak{m}^*)$ so that

$$d\omega + \theta \wedge \omega = \tfrac{1}{2}\, T(\omega \wedge \omega). \tag{0.2}$$

The function T is the *torsion function* of θ, and θ is *torsion-free* if $T = 0$. Any connection θ' on B is $\theta + p(\omega)$ where $p : B \to \mathrm{Hom}(\mathfrak{m}, \mathfrak{h})$ is an H-equivariant function. Its torsion function is $T' = T + \delta(p)$, where $\delta : \mathfrak{h} \otimes \mathfrak{m}^* \to \mathfrak{m} \otimes \Lambda^2(\mathfrak{m}^*)$ is defined via the inclusion $\mathfrak{h} \subset \mathfrak{m} \otimes \mathfrak{m}^*$ and the skewsymmetrizing map $\mathfrak{m} \otimes \mathfrak{m}^* \otimes \mathfrak{m}^* \to \mathfrak{m} \otimes \Lambda^2(\mathfrak{m}^*)$.

Evidently, the reduced map $\bar{T} : B \to \mathrm{coker}(\delta)$ is independent of the choice of θ and is the obstruction to choosing a torsion-free connection on B. For example,

when $H = \mathrm{Sp}(S)$ as above, it is not difficult to see that $\mathrm{coker}(\delta) \simeq \Lambda^3(\mathfrak{m}^*)$ and that the reduced torsion represents the exterior derivative $d\sigma$ of the associated almost symplectic form $\sigma \in \Omega^2(M)$. Thus, an $\mathrm{Sp}(S)$-structure admits a torsion-free connection if and only if the almost symplectic structure on M is actually symplectic. Furthermore, $\ker(\delta) \simeq \mathrm{S}^3(\mathfrak{m}^*)$, so that the torsion function T does not determine a unique connection θ. By contrast, when $H = \mathrm{O}(Q)$, the map δ is an isomorphism, which is merely the fundamental lemma of Riemannian geometry: Every pseudo-Riemannian metric on M possesses a unique compatible, torsion-free connection.

The reduced torsion is the first order local invariant for H-structures on M and H-structures satisfying $\bar{T} = 0$ are usually referred to as *torsion-free* or 1-*flat*. Many (but not all) of the H-structures that have received the most attention in differential geometry are 1-flat. All pseudo-Riemannian metric structures are 1-flat, an almost symplectic structure is 1-flat if and only if it is symplectic (Darboux' Theorem), an almost complex structure is 1-flat if and only if it is complex (Newlander-Nirenberg Theorem), an Hermitian structure is 1-flat if and only if it is Kähler. Contact structures and Carnot-Carathéodory structures, on the other hand, are definitely not 1-flat, as the nondegeneracy of the reduced torsion is an essential part of the geometry.

1.4. CRITERIA FOR HOLONOMY IN THE TORSION-FREE CASE. The condition $\bar{T} = 0$ is a $\mathrm{Diff}(M)$-invariant, first order equation for H-structures on M. Given a torsion-free H-structure $B \subset F(M, \mathfrak{m})$, one can ask about the possibilities for the holonomy group of a torsion-free connection θ on B. When $\ker(\delta) = 0$, a very common situation, the torsion-free connection, if it exists, will be unique, so that it makes sense to speak of the holonomy of B itself.

The vanishing of the torsion implies nontrivial restrictions on the holonomy. Assuming $T = 0$ and differentiating (0.2) yields $\Theta \wedge \omega = 0$. Thus, the *curvature function* $R : B \to \mathfrak{h} \otimes \Lambda^2(\mathfrak{m}^*)$ of θ, for which the *second structure equation of Cartan*

$$\Theta = d\theta + \tfrac{1}{2}[\theta, \theta] = \tfrac{1}{2} R(\omega \wedge \omega) \tag{0.3}$$

holds, takes values in the kernel $K(\mathfrak{h})$ of the map $\delta : \mathfrak{h} \otimes \Lambda^2(\mathfrak{m}^*) \to \mathfrak{m} \otimes \Lambda^3(\mathfrak{m}^*)$ defined by the same methods[1] as the previous δ. In particular, if there is a proper subalgebra $\mathfrak{g} \subset \mathfrak{h}$ so that $K(\mathfrak{h}) \subseteq \mathfrak{g} \otimes \Lambda^2(\mathfrak{m}^*)$, then, by the Ambrose-Singer holonomy

[1] The maps I am denoting by δ (as well as the ones to follow) are part of the *Spencer complex* associated to the inclusion $\mathfrak{h} \subset \mathfrak{m} \otimes \mathfrak{m}^*$, see [Br3]. These maps are $\mathfrak{h}$-module maps and so the various kernels and cokernels to be introduced are $\mathfrak{h}$-modules as well.

theorem, the (restricted) holonomy of θ must lie in the connected Lie group $G \subset H$ whose Lie algebra is $\mathfrak{g}$.

The intersection of the subspaces $\mathfrak{s} \subset \mathfrak{h}$ that satisfy $K(\mathfrak{h}) \subseteq \mathfrak{s} \otimes \Lambda^2(\mathfrak{m}^*)$ is an ideal $\mathfrak{g} \subset \mathfrak{h}$. Thus, a necessary condition that there exist a torsion-free connection with holonomy $H \subset \mathrm{GL}(\mathfrak{m})$ is that $K(\mathfrak{h}) \neq K(\mathfrak{g})$ for any proper ideal $\mathfrak{g} \subset \mathfrak{h}$. This is usually referred to as *Berger's first criterion* [Ber, Br3].

This criterion is very restrictive: If $\mathfrak{h}$ is semi-simple, there are, up to equivalence, only a finite number of representations $\mathfrak{h} \hookrightarrow \mathfrak{gl}(\mathfrak{m})$ without trivial summands satisfying it. As a simple example, if $\mathfrak{h} \simeq \mathfrak{sl}(2,\mathbb{R})$, and $V_k \simeq \mathsf{S}^k(V_1)$ denotes the irreducible $\mathfrak{sl}(2,\mathbb{R})$-representation of dimension $k+1$, then the $\mathfrak{sl}(2,\mathbb{R})$-representations $\mathfrak{m}$ without V_0-summands that satisfy Berger's first criterion are V_1, $V_1 \oplus V_1$, V_2, V_3, and V_4.

Symmetric examples. One large class of examples where Berger's first criterion is satisfied is provided by the following construction: Suppose that there is a surjective skewsymmetric pairing $\mathfrak{m} \times \mathfrak{m} \to \mathfrak{h}$ so that, together with the Lie algebra bracket on $\mathfrak{h}$ and the $\mathfrak{h}$-module pairing $\mathfrak{h} \times \mathfrak{m} \to \mathfrak{m}$, it defines a Lie algebra on $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$. Then the pair $(\mathfrak{g}, \mathfrak{h})$ is a symmetric pair of Lie algebras [KN]. Let G be the simply connected Lie group with Lie algebra $\mathfrak{g}$ and let $\tilde{H} \subset G$ be the (necessarily closed) connected subgroup corresponding to the subalgebra $\mathfrak{h} \subset \mathfrak{g}$ and let $H = \mathrm{Ad}_{\mathfrak{m}}(\tilde{H}) \subset \mathrm{GL}(\mathfrak{m})$ be its almost faithful image. Then $M = G/\tilde{H}$ is an affine symmetric space in a canonical way, and the coset projection $G \to M$ covers a torsion-free H-structure on M with connection whose holonomy is H (by the assumption $[\mathfrak{m}, \mathfrak{m}] = \mathfrak{h}$ and the Ambrose-Singer holonomy theorem.) The classification of the symmetric Lie algebra pairs $(\mathfrak{g}, \mathfrak{h})$ is a (rather involved) algebra problem. It was solved by Berger [Ber2] in the case that $\mathfrak{g}$ is semi-simple or when $\mathfrak{h}$ acts irreducibly on $\mathfrak{m} \simeq \mathfrak{g}/\mathfrak{h}$.

The case where θ is a locally symmetric connection with holonomy H is characterized on the H-structure B by the condition that $R : B \to K(\mathfrak{h})$ be constant. In fact, differentiating (0.3) yields $0 = (dR + \theta.R)(\omega \wedge \omega)$ (where $\theta.R$ is the result of the $\mathfrak{h}$-module pairing $\mathfrak{h} \times K(\mathfrak{h}) \to K(\mathfrak{h})$). Equivalently, $dR = -\theta.R + R'(\omega)$ where $R' : B \to K(\mathfrak{h}) \otimes \mathfrak{m}^*$ must take values in the kernel $K^1(\mathfrak{h})$ of the natural map $\delta : K(\mathfrak{h}) \otimes \mathfrak{m}^* \to \mathfrak{h} \otimes \Lambda^3(\mathfrak{m}^*)$. Now, R' vanishes identically exactly when R is parallel, which is exactly when the pair (B, θ) defines a locally symmetric affine structure on M. Thus, if H can be the holonomy of a torsion-free affine connection that is not locally symmetric, then $K^1(\mathfrak{h}) \neq 0$. This is *Berger's second criterion.*

For example, when $\mathfrak{h} = \mathfrak{sl}(2,\mathbb{R})$ and $\mathfrak{m} = V_4$, one has $K^1(\mathfrak{h}) = 0$. Thus, this 5-dimensional representation of $\mathrm{SL}(2,\mathbb{R})$ could occur as holonomy of a torsion-free connection on M^5 only when that connection is locally symmetric. In fact, it occurs

as the holonomy of the symmetric spaces $\mathrm{SL}(3,\mathbb{R})/\mathrm{SO}(2,1)$ and $\mathrm{SU}(2,1)/\mathrm{SO}(2,1)$ and in no other way. The other four cases with $\mathfrak{h} = \mathfrak{sl}(2,\mathbb{R})$ satisfying Berger's first criterion also satisfy Berger's second criterion. Moreover, each does occur as the holonomy of a torsion-free connection that is not locally symmetric.

1.5. CLASSIFICATION. In the case where $\mathfrak{m}$ is an irreducible $\mathfrak{h}$-module (implying that $\mathfrak{h}$ is reductive), Berger's fundamental works [Ber1, Ber2] went a long way towards classifying the subalgebras $\mathfrak{h} \subset \mathfrak{gl}(\mathfrak{m})$ that satisfy his first and second criteria. The methods involved heavy use of representation theory and ultimately reduced to a painstaking, elaborate case analysis. This list was refined and completed by the combined work of several people: Alekseevskii [Al1], myself, and most recently and importantly, the combined work of Chi, Merkulov, and Schwachhöfer. I will discuss this further in §3.

Berger's work provided a (partial) list of candidates for the irreducibly acting holonomy groups of torsion-free connections that are not locally symmetric. He divided the list into two parts: The first part consists of $\mathfrak{h} \subset \mathfrak{gl}(\mathfrak{m})$ that lie in some $\mathfrak{so}(Q)$ for some nondegenerate quadratic form Q on $\mathfrak{m}$, so that the associated H-structure defines a pseudo-Riemannian structure on M. For these cases, the injectivity of the map $\delta : \mathfrak{h} \otimes \mathfrak{m}^* \to \mathfrak{m} \otimes \Lambda^2(\mathfrak{m}^*)$ implies that the torsion-free connection is unique, so that it makes sense to speak of the holonomy of the underlying H-structure itself. I will refer to this part as the *metric list*. The second part consists of $\mathfrak{h} \subset \mathfrak{gl}(\mathfrak{m})$ that do not lie in any $\mathfrak{so}(Q)$ and hence will be referred to as the *nonmetric list*. For many of the subalgebras on the nonmetric list, the map δ is not injective, so that the associated H-structure does not determine the torsion-free connection.

These two lists are Tables I and II, essentially. I have taken the liberty of modifying Berger's lists slightly, dropping the entries in the original list that did not actually satisfy Berger's two criteria ($\mathrm{SO}^*(2n) \simeq \mathrm{SO}(n,\mathbb{H})$, which does not satisfy the first criterion[2], and the $\mathrm{Spin}(9,\mathbb{C})$-type entries, which do not satisfy the second criterion[3]) and including $\mathrm{Sp}(p,\mathbb{R})\cdot\mathrm{SL}(2,\mathbb{R})$, the inadvertently omitted 'split form' of the quaternionic Kähler case. In the nonmetric case, I have amplified the list somewhat by making explicit the various real forms as well as the fact that, except for $\mathrm{CO}(p,q)$, each of the entries on Berger's nonmetric list represents only the semi-simple part of H, to which one can add an arbitrary subgroup of the (abelian) commuting subgroup to make the full group H.

[2] An observation due to R. McLean.

[3] Observed independently by Alekseevskii [Al1] and Brown and Gray [BG].

I. Pseudo-Riemannian, Irreducible Holonomies in $\mathbb{R}^n$

n	H	Local Generality*
$p+q \geq 2$	$\mathrm{SO}(p,q)$	$\frac{1}{2}n(n-1)$ of n
$2p$	$\mathrm{SO}(p,\mathbb{C})$	$\frac{1}{2}p(p-1)^{\mathbb{C}}$ of $p^{\mathbb{C}}$
$2(p+q) \geq 4$	$\mathrm{U}(p,q)$	1 of n
$2(p+q) \geq 4$	$\mathrm{SU}(p,q)$	2 of $n-1$
$4(p+q) \geq 8$	$\mathrm{Sp}(p,q)$	$2(p+q)$ of $(2p+2q+1)$
$4(p+q) \geq 8$	$\mathrm{Sp}(p,q)\cdot\mathrm{Sp}(1)$	$2(p+q)$ of $(2p+2q+1)$
$4p \geq 8$	$\mathrm{Sp}(p,\mathbb{R})\cdot\mathrm{SL}(2,\mathbb{R})$	$2p$ of (2p+1)
$8p \geq 16$	$\mathrm{Sp}(p,\mathbb{C})\cdot\mathrm{SL}(2,\mathbb{C})$	$2p^{\mathbb{C}}$ of $(2p+1)^{\mathbb{C}}$
7	G_2	6 of 6
7	G_2'	6 of 6
14	$G_2^{\mathbb{C}}$	$6^{\mathbb{C}}$ of $6^{\mathbb{C}}$
8	$\mathrm{Spin}(7)$	12 of 7
8	$\mathrm{Spin}(4,3)$	12 of 7
16	$\mathrm{Spin}(7,\mathbb{C})$	$12^{\mathbb{C}}$ of $7^{\mathbb{C}}$

*Counted modulo diffeomorphism. The notation "d of ℓ" means "d functions of ℓ variables" and a superscript $\mathbb{C}$ means 'holomorphic'.

Exotic Holonomies. The nonmetric list supplied by Berger was never claimed to be a complete list, though it was supposed to have omitted at most a finite number of possibilities. The full list of omissions, comprising Tables III and IV and nowadays referred to as the *exotic list*, was recently compiled by a combination of the efforts of Chi, Merkulov, Schwachhöfer, and myself. This will be reported on in §3, along with the reasons for the division into two lists.

1.6. LOCAL EXISTENCE. Berger's lists (suitably modified) provide possibilities for irreducibly acting holonomy groups, but to verify that these possibilities actually can occur requires methods beyond representation theory. Most of the methods that have been employed can be grouped into a small number of categories:

Explicit construction. This is the simplest method, when it is available. For the metric list, there are locally symmetric examples with every holonomy except the special Kähler cases, where the holonomy is $\mathrm{SU}(p,q) \subset \mathrm{GL}(\mathbb{C}^{p+q})$; the hyper-Kähler cases, where the holonomy is $\mathrm{Sp}(p,q) \subset \mathrm{GL}(\mathbb{H}^{p+q})$; and the 'exceptional' holonomies, which comprise the groups $\mathrm{G}_2^{\mathbb{C}} \subset \mathrm{GL}(\mathbb{C}^7)$ and $\mathrm{Spin}(7,\mathbb{C}) \subset \mathrm{GL}(\mathbb{C}^8)$ and certain of their real forms. Of course, one would like to know that the locally symmetric examples are not the only ones.

II. The 'Classical' Non-Metric, Irreducible Holonomies

$H^{\diamond}$	$\mathfrak{m}$	Restrictions$^{\triangledown}$
$G_{\mathbb{R}}\cdot\mathrm{SL}(n,\mathbb{R})$	$\mathbb{R}^n$	$n\geq 2$
$G_{\mathbb{C}}\cdot\mathrm{SL}(n,\mathbb{C})$	$\mathbb{C}^n$	$n\geq 1$
$G_{\mathbb{R}}\cdot\mathrm{SL}(n,\mathbb{H})$	$\mathbb{H}^n$	$n\geq 1$
$\mathrm{Sp}(n,\mathbb{R})$	$\mathbb{R}^{2n}$	$n\geq 2$
$\mathrm{Sp}(n,\mathbb{C})$	$\mathbb{C}^{2n}$	$n\geq 2$
$\mathbb{R}^+\cdot\mathrm{Sp}(2,\mathbb{R})$	$\mathbb{R}^4$	
$\mathbb{C}^*\cdot\mathrm{Sp}(2,\mathbb{C})$	$\mathbb{C}^4$	
$\mathrm{CO}(p,q)$	$\mathbb{R}^{p+q}$	$p+q\geq 3$
$G_{\mathbb{C}}\cdot\mathrm{SO}(n,\mathbb{C})$	$\mathbb{C}^n$	$n\geq 3$
$G_{\mathbb{R}}\cdot\mathrm{SL}(p,\mathbb{R})\cdot\mathrm{SL}(q,\mathbb{R})$	$\mathbb{R}^{pq}\simeq\mathbb{R}^p\otimes_{\mathbb{R}}\mathbb{R}^q$	$p\geq q\geq 2,\ (p,q)\neq(2,2)$
$G_{\mathbb{C}}\cdot\mathrm{SL}(p,\mathbb{C})\cdot\mathrm{SL}(q,\mathbb{C})$	$\mathbb{C}^{pq}\simeq\mathbb{C}^p\otimes_{\mathbb{C}}\mathbb{C}^q$	$p\geq q\geq 2,\ (p,q)\neq(2,2)$
$G_{\mathbb{R}}\cdot\mathrm{SL}(p,\mathbb{H})\cdot\mathrm{SL}(q,\mathbb{H})$	$\mathbb{R}^{4pq}\simeq\mathbb{H}^p\otimes_{\mathbb{H}}\mathbb{H}^q$	$p\geq q\geq 1,\ (p,q)\neq(1,1)$
$G_{\mathbb{R}}\cdot\mathrm{SL}(p,\mathbb{C})$	$\mathbb{R}^{p^2}\simeq\left(\mathbb{C}^p\otimes_{\mathbb{C}}\overline{\mathbb{C}^p}\right)^{\mathbb{R}}$	$p\geq 3$
$G_{\mathbb{R}}\cdot\mathrm{SL}(p,\mathbb{R})$	$\mathbb{R}^{p(p+1)/2}\simeq\mathsf{S}^2_{\mathbb{R}}(\mathbb{R}^p)$	$p\geq 3$
$G_{\mathbb{C}}\cdot\mathrm{SL}(p,\mathbb{C})$	$\mathbb{C}^{p(p+1)/2}\simeq\mathsf{S}^2_{\mathbb{C}}(\mathbb{C}^p)$	$p\geq 3$
$G_{\mathbb{R}}\cdot\mathrm{SL}(p,\mathbb{H})$	$\mathbb{R}^{p(2p+1)}\simeq\mathsf{S}^2_{\mathbb{H}}(\mathbb{H}^p)$	$p\geq 2$
$G_{\mathbb{R}}\cdot\mathrm{SL}(p,\mathbb{R})$	$\mathbb{R}^{p(p-1)/2}\simeq\Lambda^2_{\mathbb{R}}(\mathbb{R}^p)$	$p\geq 5$
$G_{\mathbb{C}}\cdot\mathrm{SL}(p,\mathbb{C})$	$\mathbb{C}^{p(p-1)/2}\simeq\Lambda^2_{\mathbb{C}}(\mathbb{C}^p)$	$p\geq 5$
$G_{\mathbb{R}}\cdot\mathrm{SL}(p,\mathbb{H})$	$\mathbb{R}^{p(2p-1)}\simeq\Lambda^2_{\mathbb{H}}(\mathbb{H}^p)$	$p\geq 3$

$^{\diamond}$ $G_{\mathbb{F}}$ is any connected subgroup of $\mathbb{F}^*$. $^{\triangledown}$ To avoid repetition or reducibility.

Sometimes constructing examples is easy: The generic pseudo-Riemannian metric has holonomy is $\mathrm{SO}(p,q)$.

In other cases, simple underlying geometric structures can be used as a starting point. For example, all complex structures are flat, and the general (pseudo-)Kähler metric can be described in the standard background complex structure by means of a (pseudo-)Kähler potential. For the generic choice of such a potential, the holonomy will be $\mathrm{U}(p,q)\subset\mathrm{SO}(2p,2q)$. Another example is the special Kähler case, where one can start with a background complex structure with a specified holomorphic volume form and then find a Kähler metric preserving this volume form by requiring that the Kähler potential satisfy a single second order, elliptic equation.

One can also find examples by looking for those with a large symmetry group. For the hyper-Kähler case, one can start with the complex symplectic structure on

III. Exotic Conformal Holonomies

$H^{\diamondsuit}$	m
$\mathbb{R}^+ \cdot \mathrm{SL}(2,\mathbb{R})$	$\mathbb{R}^4 \simeq S^3(\mathbb{R}^2)$
$\mathbb{C}^* \cdot \mathrm{SL}(2,\mathbb{C})$	$\mathbb{C}^4 \simeq S^3(\mathbb{C}^2)$
$G_{\mathbb{C}} \cdot \mathrm{SL}(2,\mathbb{R})$ [a]	$\mathbb{C}^2 \simeq \mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}^2$
$G_{\mathbb{C}} \cdot \mathrm{Sp}(1)$ [b]	$\mathbb{C}^2 \simeq \mathbb{C} \otimes_{\mathbb{C}} \mathbb{H}$
$G_{\mathbb{R}} \cdot \mathrm{Spin}(5,5)$	$\mathbb{R}^{16}$
$G_{\mathbb{R}} \cdot \mathrm{Spin}(1,9)$	$\mathbb{R}^{16}$
$G_{\mathbb{C}} \cdot \mathrm{Spin}(10,\mathbb{C})$	$\mathbb{C}^{16}$
$G_{\mathbb{R}} \cdot \mathrm{E}_6^1$	$\mathbb{R}^{27}$
$G_{\mathbb{R}} \cdot \mathrm{E}_6^4$	$\mathbb{R}^{27}$
$G_{\mathbb{C}} \cdot \mathrm{E}_6^{\mathbb{C}}$	$\mathbb{C}^{27}$

$^{\diamondsuit}$ $G_{\mathbb{F}}$ is any connected subgroup of $\mathbb{F}^*$

Restrictions:

[a] $G_{\mathbb{C}} \not\subseteq \mathbb{R}^*$ (for irreducibility).

[b] $G_{\mathbb{C}} \not\subseteq S^1$ (to be nonmetric).

[c] $p+q \geq 3$ (for irreducibility).

[d] $n \geq 3$ (for irreducibility).

[e] $n \geq 2$ (to be nonmetric).

IV. Exotic Symplectic Holonomies

H	m
$\mathrm{SL}(2,\mathbb{R})$	$\mathbb{R}^4 \simeq S^3(\mathbb{R}^2)$
$\mathrm{SL}(2,\mathbb{C})$	$\mathbb{C}^4 \simeq S^3(\mathbb{C}^2)$
$\mathrm{SL}(2,\mathbb{R}) \cdot \mathrm{SO}(p,q)$ [c]	$\mathbb{R}^2 \otimes \mathbb{R}^{p+q}$
$\mathrm{SL}(2,\mathbb{C}) \cdot \mathrm{SO}(n,\mathbb{C})$ [d]	$\mathbb{C}^2 \otimes \mathbb{C}^n$
$\mathrm{Sp}(1) \cdot \mathrm{SO}(n,\mathbb{H})$ [e]	$\mathbb{H}^n$
$\mathrm{Sp}(3,\mathbb{R})$	$\mathbb{R}^{14} \simeq \Lambda^3_0(\mathbb{R}^6)$
$\mathrm{Sp}(3,\mathbb{C})$	$\mathbb{C}^{14} \simeq \Lambda^3_0(\mathbb{C}^6)$
$\mathrm{SL}(6,\mathbb{R})$	$\mathbb{R}^{20} \simeq \Lambda^3(\mathbb{R}^6)$
$\mathrm{SU}(1,5)$	$\mathbb{R}^{20} \simeq \Lambda^3(\mathbb{C}^6)^{\mathbb{R}}$
$\mathrm{SU}(3,3)$	$\mathbb{R}^{20} \simeq \Lambda^3(\mathbb{C}^6)^{\mathbb{R}}$
$\mathrm{SL}(6,\mathbb{C})$	$\mathbb{C}^{20} \simeq \Lambda^3(\mathbb{C}^6)$
$\mathrm{Spin}(2,10)$	$\mathbb{R}^{32}$
$\mathrm{Spin}(6,6)$	$\mathbb{R}^{32}$
$\mathrm{Spin}(12,\mathbb{C})$	$\mathbb{C}^{32}$
E_7^5	$\mathbb{R}^{56}$
E_7^7	$\mathbb{R}^{56}$
$\mathrm{E}_7^{\mathbb{C}}$	$\mathbb{C}^{56}$

the cotangent bundle of certain Hermitian symmetric spaces and look for a Kähler potential compatible with this complex symplectic structure that is invariant under the action of the isometry group, thereby reducing the problem to solving an ordinary differential equation. This was Calabi's method for constructing a hyper-Kähler metric on $T^*\mathbb{CP}^n$, the first known example in general dimensions. In the quaternionic Kähler case, where the holonomy is $\mathrm{Sp}(p,q)\cdot\mathrm{Sp}(1) \subset \mathrm{GL}(\mathbb{R}^{4(p+q)})$, Alekseevskii found homogeneous nonsymmetric examples on certain solvable Lie groups. Even for the exceptional holonomies, there are explicit examples of cohomogeneity one [Br1], [BS].

Examples in the hyper-Kähler and quaternionic Kähler cases can also be constructed by the method of *reduction*, which takes advantage of descriptions of these structures in terms of multi-symplectic geometry, generalizing the well-known method of Marsden-Weinstein reduction in symplectic geometry so as to handle the multi-symplectic case. For an account, see [Bes, Addendum E].

Twistor Methods. After Penrose's description of the self-dual metrics in dimension 4, Hitchin and Salamon [Sa], among others, were able to generalize this method to describe the hyper-Kähler and quaternionic Kähler metrics in terms of natural holomorphic geometric structures on the moduli space of rational curves with certain simple normal bundles in a complex manifold.

In fact, it was the study [Br2] of the moduli space of rational curves on a complex surface with normal bundle $\mathcal{O}(3)$ that turned up the first known examples (the first two entries in each of Tables III and IV) of omissions[4] from Berger's nonmetric list. Moreover, in the holomorphic category, it was shown that any connection with one of these holonomies could be constructed as the natural connection on the four-dimensional component of the moduli space of Legendrian rational curves in a complex contact three-fold.

Following this, Merkulov [Me1] showed that this approach could be generalized to cover the geometry of the moduli space of Legendrian deformations of certain complex homogeneous spaces and began to discover more exotic examples. This and its further developments will be reported on in §3.

Exterior differential systems. Another approach is to treat the equation $\bar{T} = 0$ directly as a system of PDE for sections of the bundle $F(M, \mathfrak{m})/H \to M$. In nearly all cases, this method leads to the study of an overdetermined system of PDE, so that Cartan-Kähler machinery must be brought to bear. For definitions and results regarding Cartan-Kähler theory, the reader can consult [BCG].

The general approach can be summarized as follows: Consider the structure equations derived so far for a torsion-free connection θ on an H-structure $B \subset F(M, \mathfrak{m})$,

$$\begin{aligned} d\omega &= -\theta \wedge \omega \\ d\theta &= -\tfrac{1}{2}[\theta, \theta] + \tfrac{1}{2} R(\omega \wedge \omega) \\ dR &= -\theta.R + R'(\omega) \end{aligned} \tag{0.4}$$

where $R : B \to K(\mathfrak{h})$ and $R' : B \to K^1(\mathfrak{h}) \subset K(\mathfrak{h}) \otimes \mathfrak{m}^*$ are as defined before. There are two things that need to be checked in order to be able to apply Cartan's general existence theorem for coframings at this level: First, the inclusion $K^1(\mathfrak{h}) \subset K(\mathfrak{h}) \otimes \mathfrak{m}^*$ should be an involutive tableau in Cartan's sense. Second, there should be a quadratic map $Q : K(\mathfrak{h}) \to K^1(\mathfrak{h}) \otimes \mathfrak{m}^*$ so that the exterior derivative of the third structure equation in (0.4) can be written in the form $\big(dR' + \theta.R' - Q(R)(\omega)\big)(\omega) = 0$. (This is the familiar 'vanishing torsion' condition in exterior differential systems.)

When these two conditions are satisfied, Cartan's existence theorem asserts that, up to local diffeomorphism, the real analytic torsion-free connections with holonomy

[4] These examples were referred to as 'exotic' in [Br2] and the term has been adopted to describe any nonmetric subgroup $H \subset \mathrm{GL}(\mathfrak{m})$ that satisfies Berger's criteria but that does not appear on Berger's original nonmetric list.

lying in $H \subset \mathrm{GL}(\mathfrak{m})$ depend on s_q functions of q variables, where s_q is the last nonzero Cartan character of $K^1(\mathfrak{h}) \subset K(\mathfrak{h}) \otimes \mathfrak{m}^*$. Moreover, for any $(R_0, R_0') \in K(\mathfrak{h}) \times K^1(\mathfrak{h})$, there exists a torsion-free connection θ on an H-structure $B \subset F(\mathfrak{m}, \mathfrak{m})$ and a $u_0 \in B$ for which the curvature functions R and R' satisfy $R(u_0) = R_0$ and $R'(u_0) = R_0'$.

When one can choose the element $R_0 \in K(\mathfrak{h}) \subset \mathfrak{h} \otimes \Lambda^2(\mathfrak{m}^*)$ so that it is surjective[5] as a map $R_0 : \Lambda^2(\mathfrak{m}) \to \mathfrak{h}$, it will then follow from the Ambrose-Singer holonomy theorem that the holonomy of θ contains the identity component of H. If, moreover, one can choose R_0' to be nonzero, such a connection θ will not be locally symmetric.

This analysis applies successfully to each of the entries of Tables I, II, and III. By contrast, for each of the entries of Table IV, the tableau $K^1(\mathfrak{h}) \subset K(\mathfrak{h}) \otimes \mathfrak{m}^*$ is not involutive. Further discussion of this point will be given in §3.

Generally, this method is good only for local analysis, but it has the distinct advantage that it not only proves existence of connections with a given holonomy, but provides their 'degree of generality' in Cartan's sense. For example, Table I gives the degree of generality of each of the possible pseudo-Riemannian, irreducible holonomies. For a similar discussion of the nonmetric list, see the survey [Br3], where various simplifications of the general argument are introduced to shorten the exposition.

This method was first used to prove the existence of metrics with holonomy G_2 and Spin(7) and is still the only method that constructs the general local solution and describes its degree of generality. This is also the only known method for analyzing Entries 3 and 4 of Table III.

Poisson Constructions. The examples H in Table IV are subgroups of $\mathrm{Sp}(S) \subset \mathrm{GL}(\mathfrak{m})$ for a nondegenerate skewsymmetric bilinear form S on $\mathfrak{m}$. Hence the corresponding H-structures (when they exist) have an underlying symplectic structure.

For the first two examples from Table IV, each torsion-free connection (M, B, θ) of these types was analyzed and reconstructed in [Br2] from its derived curvature map $J = (R, R') : B \to K(\mathfrak{h}) \oplus K^1(\mathfrak{h})$. This reconstruction involved a number of seemingly miraculous identities, but since only these two examples were known, it did not seem worthwhile to look for an interpretation of these identities. However, when Chi, Merkulov, and Schwachhöfer [CMS] found other exotic symplectic examples, they noticed that this reconstruction technique generalized and they were able to explain it in the context of Poisson geometry in a very beautiful way. This, too, will be reported on in §3.

[5] Actually, one only needs that the image $R_0\big(\Lambda^2(\mathfrak{m})\big) \subset \mathfrak{h}$ generates $\mathfrak{h}$ as an algebra.

1.7. COMPACT RIEMANNIAN EXAMPLES. The history of compact Riemannian manifolds with reduced holonomy groups is long and complex, so I will not attempt a full account here. For more details on the cases I only mention, the reader can consult the relevant chapters of [Bes] and the references cited therein.

Kähler manifolds. This subject has the longest history, predating even Berger's classification. Every smooth algebraic variety carries a Kähler structure and this accounts for their importance in algebraic geometry.

Special Kähler manifolds. The major milestone here is, of course, Yau's solution in the mid 1970s of the Calabi conjecture, showing that every compact Kähler manifold with trivial canonical bundle carries a special Kähler structure. For this reason, compact manifolds endowed with such a structure are usually referred to as Calabi-Yau manifolds.

Hyper-Kähler manifolds. In the early 1980s, Beauville and Mukai were each able to construct compact, simply connected Kähler manifolds M^{4p} that carried a nondegenerate complex symplectic form (Fujiki had constructed examples for $p = 2$ slightly earlier). These necessarily had trivial canonical bundle, so by Yau's solution of the Calabi conjecture, these carried special Kähler structures. By an argument of Bochner, the complex symplectic form had to be parallel with respect to this special Kähler structure, which forced the holonomy to lie in $\mathrm{Sp}(p)$. Further arguments showed that these examples were not products of lower dimensional complex manifolds and this implied that the holonomy had to actually be $\mathrm{Sp}(p)$. These were the first known compact examples.

Quaternion Kähler manifolds. In this case, all known compact examples are locally symmetric, but we know of no reason why this should be true, except for dimension 8 [PoS]. Of course, a great deal is known about the possible geometry of such examples, see [Sa].

G_2 *and* $\mathrm{Spin}(7)$ *manifolds.* The remarkable recent work of Joyce [Jo1,2] establishes the existence of compact manifolds with these holonomies. This will be reported on in §2.

2. RIEMANNIAN HOLONOMY

In this section, g will denote a smooth Riemannian metric on a connected, smooth manifold M^n. The reference space $\mathfrak{m}$ will be taken to be $\mathbb{R}^n$ with its standard inner product, and $\mathrm{O}(\mathfrak{m}) \subset \mathrm{GL}(\mathfrak{m})$ will denote its orthogonal group. The $\mathrm{O}(\mathfrak{m})$-structure $B \subset F(M, \mathfrak{m})$ consisting of the coframes $u : T_xM \to \mathfrak{m}$ that are isometries

of vector spaces is the orthonormal coframe bundle and the Levi-Civita connection on B will be denoted θ. This is, of course, the unique torsion-free $\mathrm{O}(\mathfrak{m})$-connection on B. When there is no danger of confusion, I will simply write H_u and B_u instead of H_u^θ and B_u^θ.

One thing that makes the Riemannian case simpler to deal with than others is the *de Rham Splitting Theorem* [Bes, KNo], which occurs in a local form and a global form. The local form asserts that if $(H_u)^0$ acts reducibly on $\mathfrak{m}$, say, preserving irreducible orthogonal subspaces $\mathfrak{m}_i \subset \mathfrak{m}$ for $1 \leq i \leq k$, then $(H_u)^0$ is the direct product of its subgroups $(H_u)_i^0$, where $(H_u)_i^0$ is the subgroup that acts trivially on $\mathfrak{m}_j$ for $j \neq i$. Moreover, the metric g locally splits as a product in a corresponding way[6]. The global form asserts that if, in addition, M is simply connected and the metric g is complete, then M globally splits as a Riemannian product

$$(M, g) = (M_1, g_1) \times \cdots \times (M_k, g_k).$$

so that $(H_u)_i^0 = (H_u)_i$ is the holonomy of (M_i, g_i).

2.1. NON-CLOSED HOLONOMY. While the remainder of this report deals only with connected holonomy groups, I cannot pass up the opportunity to mention a recent result of particular interest in Riemannian holonomy. It had been a question for some time whether the holonomy of a compact Riemannian manifold must necessarily be compact. Using the de Rham Splitting Theorem and Berger's holonomy classification in the irreducible Riemannian case, one sees that this is so if M is simply connected. Thus, for any Riemannian manifold, the restricted holonomy group is compact, so it becomes a question of whether the fundamental group can cause the holonomy group to have an infinite number of components even when M is compact.

Very recently, B. Wilking [Wi] has shown that this can indeed occur. He has produced an example of a compact manifold whose holonomy group does have an infinite number of components. His example is of the form $M^5 = \Gamma\backslash(\mathbb{R}^2 \times N^3)$ where Γ is a subgroup of the isometry group of $\mathbb{R}^2 \times N^3$ that acts properly discontinuously and cocompactly.

2.2. COMPACT MANIFOLDS WITH EXCEPTIONAL HOLONOMY. The most remarkable development in Riemannian holonomy in recent years has been the spectacular con-

[6] What is also true, but not obvious, is that each of the groups $(H_u)_i^0 \subset \mathrm{SO}(\mathfrak{m}_i)$ is the holonomy of some Riemannian metric, even when it is not the holonomy of the corresponding local factor of g. See [Bes, Theorem 10.108]

struction by Dominic Joyce of compact 7-manifolds with holonomy G_2 and compact 8-manifolds with holonomy Spin(7). His constructions are full of new ideas and, while it is not difficult to outline these ideas, their successful execution turns out to require very careful, subtle estimates. I will not attempt to explain these, but refer the reader to the original sources [Jo1,2] and to the forthcoming book [Jo3]. Also, I will concentrate on the G_2 case, as the Spin(7) case follows the same spirit, but the details are different.

The fundamental 3-form. Let $\mathfrak{m} = \mathbb{R}^7$. It has been known for some time [Br1] that there is an open $GL(\mathfrak{m})$-orbit $\Lambda^3_+(\mathfrak{m}^*)$ in the 3-forms on $\mathfrak{m}$ so that the stabilizer of any element $\phi \in \Lambda^3_+(\mathfrak{m}^*)$ is a compact connected simple Lie group of dimension 14 and which is therefore isomorphic to G_2. Consequently, for any 7-manifold M, there is an open subbundle $\Lambda^3_+(T^*M) \subset \Lambda^3(T^*M)$ so that the G_2-structures on M are in one-to-one correspondence with the sections $\Omega^3_+(TM)$ of this bundle. Such a section will exist if and only if M is orientable and spinnable [LM], so I assume this from now on.

Associated to any section $\sigma \in \Omega^3_+(M)$, there is a canonical metric g_σ and orientation, inducing a well-defined Hodge star operator $*_\sigma : \Omega^p(M) \to \Omega^{7-p}(M)$. If σ is parallel with respect to the Levi-Civita connection of g_σ then the holonomy of g_σ will be a subgroup of G_2 and the G_2-structure associated to σ will be torsion-free. By a result of Gray, this G_2-structure is torsion-free if and only if both σ and $*_\sigma\sigma$ are closed. Conversely, if a Riemannian metric g on M^7 has holonomy a subgroup of G_2, then there will be a g-parallel section $\sigma \in \Omega^3_+(M)$ for which $g = g_\sigma$.

A metric with holonomy a subgroup of G_2 is known to be Ricci flat [Bes, 10.64], so if M is compact, Bochner vanishing shows that its harmonic 1-forms and its Killing fields must be parallel. Thus, in the compact case, $b_1(M) = 0$ if the holonomy actually equals all of G_2. Combining this information with the Cheeger-Gromoll splitting theorem [Bes, 6.67], one finds that a compact manifold with holonomy G_2 must have finite fundamental group, so one can, without loss of generality, assume that M is simply connected, which I do from now on. Conversely, a compact simply connected Riemannian 7-manifold whose holonomy is a subgroup of G_2 must actually equal G_2 since any proper subgroup of G_2 that can be a holonomy group fixes a nontrivial subspace.

One can now consider the moduli space $\mathcal{M}$ consisting of the sections $\sigma \in \Omega^3_+(M)$ satisfying $d\sigma = d*_\sigma\sigma = 0$ modulo diffeomorphisms of M isotopic to the identity. There is an obvious 'Torelli' map $\tau : \mathcal{M} \to H^3_{\mathrm{dR}}(M)$ defined by $\tau\big([\sigma]\big) = [\sigma]_{\mathrm{dR}}$. Joyce's first striking result [Jo1] is the analog of the local Torelli theorem, namely

that τ is locally one-to-one and onto, in fact, a local diffeomorphism in the natural smooth structure on $\mathcal{M}$. Thus, the moduli space is said to be 'unobstructed'.

Next, Joyce proves a remarkable existence theorem: If $\sigma \in \Omega^3_+(M)$ is closed, then there is a constant C that depends on the norm of the curvature of the metric g_σ, its volume, and its injectivity radius, so that, if $\left|d(*_\sigma\sigma)\right|_\sigma < C$, then there exists an exact 3-form ϕ so that $\sigma+\phi$ lies in $\Omega^3_+(M)$ and is closed and coclosed. In other words, a closed 3-form in $\Omega^3_+(M)$ that is 'close enough' to being coclosed can be perturbed to a cohomologous 3-form in $\Omega^3_+(M)$ that is both closed and coclosed.

Thus, to prove the existence of a compact Riemannian 7-manifold with holonomy G_2, it suffices to construct a simply connected 7-manifold endowed with a closed 3-form that satisfies such a 'close enough' estimate.

Joyce's idea for doing this is extremely clever: He starts with the flat G_2-structure σ_0 on the 7-torus $T^7 = \mathbb{R}^7/\mathbb{Z}^7$ and passes to a simply connected quotient orbifold $X = \Gamma\backslash T$ where Γ is a finite group of σ_0-symmetries. This provides a flat G_2-orbifold whose singular locus is a finite number of 3-tori T^3, each of which has a neighborhood of the form $T^3 \times B^4/\{\pm\}$ where $B^4/\{\pm\}$ is the standard 4-ball around the origin in $\mathbb{R}^4 = \mathbb{C}^2$ divided by the equivalence relation $v \sim -v$.

Now, it has been known for a long time that $\mathbb{R}^4/\{\pm\}$ is metrically the scaling limit of the SU(2)-holonomy metric on $T^*\mathbb{CP}^1$ constructed by Eguchi and Hanson as one scales the metric to contract the zero section to a point. Because $I_3\times\mathrm{SU}(2) \subset \mathrm{SO}(7)$ is a subgroup of G_2, it follows that one can regard the flat G_2-structure on X in each singular locus neighborhood $T^3 \times B^4/\{\pm\}$ as the limit of a scaling of a G_2-structure on $T^3\times T^*\mathbb{CP}^1$. Joyce cuts out these singular neighborhoods and glues back in $T^3 \times N$ where N is a neighborhood of the zero section in $T^*\mathbb{CP}^1$, smoothly joining the flat 3-form with the Eguchi-Hanson-derived (closed) 3-form on the overlaps. This 'surgery' produces a smooth manifold $\hat{X}$, but does not disturb the fundamental group, which remains trivial.

By being very careful (here is where Joyce's estimates are extremely delicate), he shows that he can do this in such a way that the resulting closed 3-form $\sigma \in \Omega^3_+(\hat{X})$ satisfies his estimate. I.e., it is close enough to being coclosed that it can be perturbed to a $\tilde{\sigma} \in \Omega^3_+(\hat{X})$ that is both closed and coclosed. Of course, since X is simply connected, it follows that the resulting torsion-free G_2-structure has holonomy equal to G_2.

By applying this idea to a number of different finite groups Γ, Joyce has been able to construct G_2-metrics on a number of different 7-manifolds.

A similar set of ideas allows Joyce to construct compact 8-manifolds with holo-

nomy Spin(7), once one realizes that Spin(7) can be defined in $\mathrm{GL}(8,\mathbb{R})$ as the stabilizer of a certain 4-form in eight variables. The interested reader should consult [Jo1,2].

3. IRREDUCIBLE TORSION-FREE NON-METRIC CONNECTIONS

3.1. TWISTOR CONSTRUCTIONS. As previously mentioned, I found the first examples of 'exotic' holonomy groups by studying the geometry of the moduli space M of rational curves (i.e., complex curves of genus 0) in a complex surface S with normal bundle $\mathcal{O}(3)$. Following the examples of Hitchin's study of rational curves in a surface S with normal bundle $\mathcal{O}(k)$ for $k = 1$ and 2, I knew that M would have dimension 4 and would have a natural G_3-structure, where $G_3 \subset \mathrm{GL}(4,\mathbb{C})$ is the image of $\mathrm{GL}(2,\mathbb{C})$ acting by linear substitutions on the 4-dimensional space V_3 of cubic polynomials in two variables. I also knew that there would be a canonical G_3-connection from general principles, but I was very surprised to find that this connection was torsion-free.

In the examples Hitchin had analyzed, the geometry on the moduli space allowed one to reconstruct the surface S and so I fully expected to be able to do the same in this case. However, it turned out that the story was more subtle than that. In the standard double fibration picture:

$$\begin{array}{ccc} & I & \\ {\scriptstyle\lambda}\swarrow & & \searrow{\scriptstyle\rho} \\ M & & S \end{array}$$

where $I \subset M \times S$ is the set of pairs (C,p) where $p \in S$ is a point of the rational curve $C \in M$ and λ and ρ are just the projections onto the factors, each $p \in S$ would correspond to a hypersurface $H_p = \lambda\big(\rho^{-1}(p)\big)$ in M (since it is only one condition for a curve in S to pass through a given point p). The members of this 2-parameter family of hypersurfaces in M would be expected to be the solutions of some differential geometric problem in M, but I was not able to find a geometrically defined 2-parameter family of hypersurfaces in the general 4-manifold carrying a torsion-free G_3-structure.

However, a 3-parameter family Y of 2-dimensional surfaces did present itself. This can be described as follows: By the defining property of a G_3-structure B on M^4, each tangent space T_xM can be thought of as the space of homogeneous cubic polynomials in two variables. This defines a distinguished $\mathbb{P}^1$ of lines, namely the perfect

cubes, and a distinguished $\mathbb{P}^1$ of 2-planes, namely the multiples of a perfect square. I called such lines and 2-planes *null*. It was not difficult to show that, when B had a torsion-free connection, each null 2-plane was tangent to a unique totally geodesic 2-surface in M all of whose tangent planes were null. This family Y then fit into a double fibration[7]

$$\begin{array}{ccc} & N^5 & \\ {\scriptstyle\lambda}\swarrow & & \searrow{\scriptstyle\rho} \\ M^4 & & Y^3 \end{array}$$

where the fibers of λ are $\mathbb{P}^1$'s. Moreover, I was able to show that Y carried a natural structure as a contact manifold, that the family of $\mathbb{P}^1$'s given by $C_x = \rho\big(\lambda^{-1}(x)\big)$ for $x \in M$ were all Legendrian curves for this contact structure, and that, moreover, this family was an open set in the space of Legendrian rational curves in Y. (In the surface case that I had started out with, Y turned out to be the projectivized tangent bundle of S.)

I then showed that the picture could be reversed: Starting with a holomorphic contact 3-manifold Y, one could look at the moduli space M of rational Legendrian curves C in Y to which the contact bundle $L \subset T^*Y$ restricts to be isomorphic to $\mathcal{O}(-3)$ and show that it was a smooth moduli space of dimension 4 on which there was a canonical torsion-free G_3-structure.

Merkulov's generalization. In a remarkable series of papers, Merkulov [Me1,2,3] showed that this moduli space and double fibration construction obtains in a very general setting, starting from the data of an irreducibly acting (and therefore reductive) complex subgroup $H \subset \mathrm{GL}(n,\mathbb{C})$, a complex n-manifold M, and a holomorphic H-structure $B \subset F(M,\mathbb{C}^n)$ endowed with a holomorphic torsion-free connection θ.

The semi-simple part $H_s \subset H$ acts irreducibly on $\mathbb{C}^n$. With respect to a Cartan subalgebra of H_s and an ordering of its roots, there will be a unique line $E \subset (\mathbb{C}^n)^*$ spanned by a vector of highest weight. The H_s-orbit $F \subset \mathbb{P}\big((\mathbb{C}^n)^*\big)$ of E is a minimal H_s-orbit in $\mathbb{P}\big((\mathbb{C}^n)^*\big)$, a generalized flag variety of H_s of some dimension $k \leq n-1$, endowed with the hyperplane section bundle $\mathcal{L}$. (In the original case I treated, F is the projectivization of the set of perfect cubes and hence is a $\mathbb{P}^1$. The bundle $\mathcal{L}$ is $\mathcal{O}(-3)$.) The H-structure B provides identifications $T_xM \simeq \mathbb{C}^n$ unique up to an action of H, so there is a subbundle $N \subset \mathbb{P}(T^*M)$ whose fiber $N_x \subset \mathbb{P}(T^*_xM)$ over x corresponds to $F \subset \mathbb{P}\big((\mathbb{C}^n)^*\big)$ via any B-identification.

[7] Assuming Y to be Hausdorff in its natural topology, which can always be arranged by replacing M by a θ-convex open set in M.

The projectivized cotangent bundle of any manifold is canonically a contact manifold, and the torsion-free condition on θ immediately implies that N is an *involutive* submanifold of $\mathbb{P}(T^*M)$, i.e., the restriction of the contact structure to N is degenerate, with Cauchy leaves of the largest possible dimension, namely $n-k-1$, the codimension of N in $\mathbb{P}(T^*M)$. When it is Hausdorff [8], the leaf space Y of this Cauchy foliation is canonically a contact manifold, yielding the double fibration

$$\begin{array}{ccc} & N^{n+k} & \\ {\scriptstyle\lambda}\swarrow & & \searrow{\scriptstyle\rho} \\ M^n & & Y^{2k+1} \end{array}$$

where the manifolds $F_x = \rho\big(\lambda^{-1}(x)\big)$ are Legendrian k-dimensional submanifolds of Y.

Merkulov then goes on to prove that, nearly always, one can recover M as the complete moduli space of the Legendrian immersions $F \subset Y$ that pull back the contact bundle $L \subset T^*Y$ to be $\mathcal{L}$. Moreover, when H_s acts as the full biholomorphism group of F (which, again, is nearly always) one can recover the original H-structure on M up to conformal scaling from the family of submanifolds $S_y = \lambda\big(\rho^{-1}(y)\big)$ for $y \in Y$.

Finally, Merkulov gives representation theoretic criteria on an irreducibly acting subgroup $H \subset \mathrm{GL}(\mathfrak{m})$ with associated generalized flag variety $F \subset \mathbb{P}(\mathfrak{m}^*)$ of dimension k and hyperplane bundle $\mathcal{L}$, which guarantee that taking a $(2k+1)$-dimensional contact manifold Y and considering the moduli space $M(Y)$ of Legendrian embeddings $F \subset Y$ that pull back the contact bundle L to be $\mathcal{L}$ yields a smooth moduli space endowed with an H-structure and a torsion-free connection.

This last step is extremely important, for it provides a way to determine which irreducibly acting subgroups $H \subset \mathrm{GL}(\mathfrak{m})$ can occur as torsion-free holonomy in terms of representation theory, specifically, in terms of the vanishing of certain H-representations constructed functorially from $\mathfrak{m}$. This provides a different approach to solving the torsion-free holonomy problem, one that was carried out successfully by a combination of efforts of Chi, Merkulov, and Schwachhöfer. In particular, this approach led to the discovery of the remaining groups in Table IV and, finally, the proof that Tables I, II, III, and IV exhaust the possibilities for irreducibly acting torsion-free holonomy.

3.2. POISSON CONSTRUCTIONS. The straightforward application of exterior differential systems to the holonomy problem outlined in §0.6 does not work for the entries

[8] This can always be arranged by replacing M by a suitably θ-convex open set in M.

of Table IV, at least at the level described so far. To see where the problem is, recall the structure equations derived so far for a torsion-free connection with holonomy $H \subset \mathrm{GL}(\mathfrak{m})$. They are

$$\begin{aligned} d\omega &= -\theta \wedge \omega \\ d\theta &= -\tfrac{1}{2}[\theta,\theta] + \tfrac{1}{2}\,R(\omega \wedge \omega) \\ dR &= -\theta.R + R'(\omega) \end{aligned} \tag{3.1}$$

where $R : B \to K(\mathfrak{h})$ and $R' : B \to K^1(\mathfrak{h}) \subset K(\mathfrak{h}) \otimes \mathfrak{m}^*$ are as already defined.

When one considers the first entry of Table IV, where $H = \mathrm{SL}(2,\mathbb{R})$ and $\mathfrak{m} \simeq V_3 = \mathsf{S}^3(V_1)$, it is not difficult to see that $K(\mathfrak{h}) \simeq V_2 = \mathsf{S}^2(V_1)$ has dimension 3 and that $K^1(\mathfrak{h}) \simeq V_3$ has dimension 4. Its Cartan characters are $(s_1, s_2, s_3, s_4) = (3, 1, 0, 0)$ but, as is easily computed, the prolongation $K^2(\mathfrak{h}) \subset K^1(\mathfrak{h}) \otimes \mathfrak{m}^*$ has dimension 1, so the tableau is not involutive and Cartan's existence theorem cannot be applied at this level.

However, there is a quadratic map $Q : K(\mathfrak{h}) \to K^1(\mathfrak{h}) \otimes \mathfrak{m}^*$ so that the exterior derivative of the third equation in (3.1) is $\big(dR' + \theta.R' - Q(R)(\omega)\big)(\omega) = 0$, implying that there is a function $R'' : B \to K^2(\mathfrak{h})$ so that the equation

$$dR' = -\theta.R' + \big(R'' + Q(R)\big)(\omega) \tag{3.2}$$

holds. Moreover, it is possible to choose the quadratic map Q in a unique way so that differentiating this last equation yields $dR''(\omega) = 0$. Now the second prolongation of $K^1(\mathfrak{h})$ vanishes, so this forces the structure equation

$$dR'' = 0. \tag{3.3}$$

Obviously, differentiating this equation will yield no new information.

At this point, Cartan's general existence theorem for coframings satisfying prescribed differential identities (a generalization of Lie's third fundamental theorem) can be applied to the entire ensemble

$$\begin{aligned} d\omega &= -\theta \wedge \omega, \\ d\theta &= -\tfrac{1}{2}[\theta,\theta] + \tfrac{1}{2}\,R(\omega \wedge \omega), \\ dR &= -\theta.R + R'(\omega), \\ dR' &= -\theta.R' + \big(R'' + Q(R)\big)(\omega), \\ dR'' &= 0. \end{aligned} \tag{3.4}$$

His theorem implies that for every $(R_0, R_0', R_0'') \in K(\mathfrak{h})\times K^1(\mathfrak{h})\times K^2(\mathfrak{h})$, there is a torsion-free H-structure $B \subset F(\mathfrak{m}, \mathfrak{m})$, unique up to local diffeomorphism, so that, at a point $u_0 \in B$ one has $R(u_0) = R_0$, $R'(u_0) = R_0'$, and $R''(u_0) = R_0''$. Consequently, the space of diffeomorphism classes of germs of such torsion-free H-structures is finite dimensional.

What Chi, Merkulov, and Schwachhöfer show in [CMS] is that this exact same picture holds for each entry in Table IV. Namely, $K(\mathfrak{h}) \simeq \mathfrak{h}$, $K^1(\mathfrak{h}) \simeq \mathfrak{m}$, $K^2(\mathfrak{h}) \simeq \mathbb{F}$ $(= \mathbb{R}$ or $\mathbb{C})$, and there is a quadratic map $Q : \mathfrak{h} \to \mathfrak{m} \otimes \mathfrak{m}^*$ so that the above analysis of the structure equations goes through exactly the same as for the original two cases (with the obvious interpretations of the pairings).

Cartan's theorem then applies and yields not only the existence of torsion-free connections with these prescribed holonomies, but that the space of germs of such H-structures, modulo diffeomorphism, is finite dimensional.

In the original two cases that I treated, the finer understanding of the moduli space of solutions entailed understanding how the images $(R, R', R'')(B) \subset \mathfrak{h}\times\mathfrak{m}\times\mathbb{F}$ partition $\mathfrak{h}\times\mathfrak{m}\times\mathbb{F}$ into subsets. This analysis would have been hopeless were it not for several (at the time) amazing identities that I found by brute force calculation. They even allowed me to prove existence without using Cartan's existence theorem.

What is shown in [CMS], however, is that these mysterious identities can be explained in terms of a natural Poisson structure on the space $\mathfrak{h}\times\mathfrak{m}\times\mathbb{F}$ (actually, they regard the last factor as a parameter and consider Poisson structures on $\mathfrak{h}\times\mathfrak{m}$). The images $(R, R', R'')(B)$ turn out to be the symplectic leaves of this Poisson structure and this point of view simplifies the reconstruction of the H-structure from the leaf image (though it does not entirely remove some of the global difficulties having to do with the symplectic realizations necessary in their construction).

3.3. ALGEBRAIC CLASSIFICATION. Once the full list of the irreducible torsion-free holonomies was known, there arose the question of whether this list could be derived through Berger's original approach, i.e., representation theory. Schwachhöfer [Schw] has shown that this can indeed be done. (As is so often the case, knowing the answer in advance helps to organize the proof.) His fully algebraic classification of the irreducibly acting subgroups $H \subset \mathrm{GL}(\mathfrak{m})$ that satisfy Berger's first criterion and the subset of those that also satisfy Berger's second criterion still involves quite a bit of case checking, but the general outline of the argument is clear.

What is particularly intriguing is Ziller's observation (reported in [Schw]) that this list can be constructed by a simple Ansatz starting from the list of Hermitian

and quaternionic symmetric spaces. A direct proof of Ziller's Ansatz would be highly desirable.

3.4. TWO LEFTOVER CASES. As of this writing, each entry in the four Tables, save two, Entries 3 and 4 in Table III, has been treated in the literature and shown to occur as holonomy, either by twistor methods or exterior differential systems methods. Existence proofs by twistor methods have some difficulty when $\mathfrak{m}$ is a complex vector space and the group $H \subset \mathrm{GL}(\mathfrak{m})$ is of the form $H = G_{\mathbb{C}} \cdot H_s$ where H_s is the semi-simple part and $G_{\mathbb{C}} \subset \mathbb{C}^*$ is a one-dimensional subgroup of $\mathbb{C}^*$, acting as scalar multiplication on $\mathfrak{m}$. The method of exterior differential systems does not have this problem, but each case does require a separate treatment.

In my survey article [Br3], I left these two entries unsettled in the case where $G_{\mathbb{C}}$ had dimension 1 because, at the time, they did not seem that interesting. Now that they are the last unsettled cases, it seems to be a good idea to resolve them, so I will do that here, though, for lack of space, I will not provide details, just give the results of the Cartan-Kähler analysis.

For the first case, $H = G_{\mathbb{C}} \cdot \mathrm{SL}(2,\mathbb{R}) \subset \mathrm{GL}(2,\mathbb{C})$, one must assume that $G_{\mathbb{C}} \not\subseteq \mathbb{R}^*$, otherwise H will not act irreducibly on $\mathfrak{m} \simeq \mathbb{C}^2$. This leaves a one parameter family of possibilities $H_\lambda = \{e^{(i+\lambda)t} \,|\, t \in \mathbb{R}\} \subset \mathbb{C}^*$ where λ is any real number. By conjugation, one can assume that $\lambda \geq 0$, so I will do this. It turns out that there are two cases:

If $\lambda = 0$, so that $H_0 = S^1$, it is not difficult to compute that $K(\mathfrak{h}) \simeq V_4 \oplus V_2 \oplus V_0$, while $K^1(\mathfrak{h}) \simeq 2V_5 \oplus 2V_3 \oplus 2V_1$ and is involutive, with characters $(s_1, s_2, s_3, s_4) = (9, 9, 5, 1)$. Moreover, the torsion is absorbable. By Cartan's theorem, solutions exist and depend on one function of four variables.

However, if $\lambda > 0$, so that $H_\lambda \simeq \mathbb{R}^+$, one computes that $K(\mathfrak{h}) \simeq V_4 \oplus V_2$ while $K^1(\mathfrak{h}) \simeq 2V_5 \oplus 2V_3$ and is involutive, with characters $(s_1, s_2, s_3, s_4) = (8, 8, 4, 0)$. Moreover, the torsion is absorbable. By Cartan's theorem, solutions exist and depend on four functions of three variables.

For the second case, $H = G_{\mathbb{C}} \cdot \mathrm{SU}(2) \subset \mathrm{GL}(2,\mathbb{C})$, one must assume that $G_{\mathbb{C}} \not\subseteq S^1$, otherwise $H = \mathrm{U}(2)$ will preserve a metric on $\mathfrak{m} \simeq \mathbb{C}^2$. This leaves a one parameter family of possibilities $J_\lambda = \{e^{(1+i\lambda)t} \,|\, t \in \mathbb{R}\} \subset \mathbb{C}^*$ where λ is any real number. By conjugation, one can assume that $\lambda \geq 0$, though I won't need to do this. Here there is only one case: One computes that $K(\mathfrak{h}) \simeq V_4^{\mathbb{R}} \oplus V_2^{\mathbb{R}} \simeq \mathbb{R}^8$ while $K^1(\mathfrak{h}) \simeq V_5 \oplus V_3 \simeq \mathbb{C}^{10} \simeq \mathbb{R}^{20}$. The tableau is involutive, with characters $(s_1, s_2, s_3, s_4) = (8, 8, 4, 0)$. Moreover, the torsion is absorbable. By Cartan's theorem, solutions exist and depend on four functions of three variables.

REFERENCES

[Al1] D. ALEKSEEVSKII - *Riemannian spaces with unusual holonomy groups*, Funct. Anal. Appl. **2** (1968), 97–105.

[Al2] D. ALEKSEEVSKII - *Classification of quaternionic spaces with a transitive solvable group of motions*, Math. USSR-Izv **9** (1975), 297–339.

[AS] A. AMBROSE and I. SINGER - *A theorem on holonomy*, Trans. Amer. Math. Soc. **75** (1953), 428–443.

[BB] L. BÉRARD-BERGERY - *On the holonomy of Lorentzian manifolds*, Proc. Symp. Pure Math. **54** (1993), 27–40.

[BI] L. BÉRARD-BERGERY and A. IKEMAKHEN - *Sur l'holonomie des variétés pseudo-riemanniennes de signature* (n, n), Bull. Soc. Math. France **125** (1997), 93–114.

[Ber1] M. BERGER, *Sur les groupes d'holonomie des variétés à connexion affine et des variétés riemanniennes*, Bull. Soc. Math. France **83** (1955), 279–330.

[Ber2] M. BERGER - *Les espaces symétriques non compacts*, Ann. Sci. École Norm. Sup. **74** (1957), 85–177.

[Bes] A. BESSE - *Einstein Manifolds*, Ergebnisse der Mathematik und ihrer Grenzgebiete, 3. Folge, Band 10, Springer-Verlag, Berlin, 1987.

[BL] A. BOREL and A. LICHNEROWICZ - *Groupes d'holonomie des variétés riemanniennes*, C.R. Acad. Sci. Paris **234** (1952), 1835–1837.

[BG] R. BROWN and A. GRAY - *Riemannian manifolds with holonomy group* Spin(9), in *Differential Geometry (in honor of Kentaro Yano)*, Kinokuniya, Tokyo, 1972, 41–59.

[Br1] R. BRYANT - *Metrics with exceptional holonomy*, Ann. Math. **126** (1987), 525–576.

[Br2] R. BRYANT - *Two exotic holonomies in dimension four, path geometries, and twistor theory*, Amer. Math. Soc. Proc. Symp. Pure Math. **53** (1991), 33–88.

[Br3] R. BRYANT - *Classical, exceptional, and exotic holonomies: a status report*, in *Actes de la Table Ronde de Géométrie Différentielle en l'Honneur de Marcel Berger*, Soc. Math. France, 1996, 93–166.

[BCG] R. BRYANT, *et al* - *Exterior Differential Systems*, MSRI Series **18**, Springer-Verlag, 1991.

[BS] R. BRYANT and S. SALAMON - *On the construction of some complete metrics with exceptional holonomy*, Duke Math. J. **58** (1989), 829–850.

[Ca1] É. CARTAN - *La Géometrie des Espaces de Riemann*, Mémorial des Sciences

Mathématiques, Fasc. IX (1925). (Especially, Chapitre VII, Section II.)

[Ca2] É. CARTAN - *Les groupes d'holonomie des espaces généralisés*, Acta. Math. **48** (1926), 1–42.

[CMS] Q.-S. CHI, S. MERKULOV, and L. SCHWACHHÖFER - *On the existence of infinite series of exotic holonomies.* Inventiones Math. **126** (1996), 391–411.

[ChS] Q.-S. CHI and L. SCHWACHHÖFER - *Exotic holonomy on moduli spaces of rational curves*, Differential Geom. App. **8** (1998), 105–134.

[Gr] M. GROMOV - *Carnot-Carathéodory spaces seen from within*, in *Sub-Riemannian geometry*, Progr. Math. **144**, Birkhäuser, Basel, 1996, 79–323.

[Jo1] D. JOYCE - *Compact Riemannian 7-manifolds with holonomy* G_2*:* I & II, J. Differential Geom. **43** (1996), 291–328, 329–375.

[Jo2] D. JOYCE - *Compact* 8*-manifolds with holonomy* Spin(7), Inventiones Math. **123** (1996), 507–552.

[Jo3] D. JOYCE - *Compact manifolds with special holonomy*, Oxford University Press, Oxford, 2000 (projected).

[KNa] S. KOBAYASHI and K. NAGANO - *On filtered Lie algebras and geometric structures,* II, J. Math. Mech. **14** (1965), 513–521.

[KNo] S. KOBAYASHI and K. NOMIZU - *Foundations of Differential Geometry, Vols.* 1 & 2, Wiley-Interscience, New York, 1963.

[LM] H. B. LAWSON, JR. and M. L. MICHELSOHN - *Spin geometry*, Princeton Math. Series **38**, Princeton Univ. Press, Princeton, NJ, 1989.

[Me1] S. MERKULOV - *Moduli spaces of compact complex submanifolds of complex fibered manifolds*, Math. Proc. Camb. Phil. Soc. **118** (1995), 71–91.

[Me2] S. MERKULOV - *Geometry of Kodaira moduli spaces*, Proc. Amer. Math. Soc. **124** (1996), 1499–1506.

[Me3] S. MERKULOV - *Existence and geometry of Legendre moduli spaces*, Math. Z. **226** (1997), 211–265.

[MS1] S. MERKULOV and L. SCHWACHHÖFER - *Classification of irreducible holonomies of torsion-free affine connections*, Ann. Math. (to appear).

[MS2] S. MERKULOV and L. SCHWACHHÖFER - *Twistor solution of the holonomy problem*, 395–402, The Geometric Universe, Science, Geometry and the work of Roger Penrose, S.A. Hugget (ed.), Oxford Univ. Press, 1998.

[PoS] Y. S. POON and S. SALAMON - *Quaternionic Kähler* 8*-manifolds with positive scalar curvature*, J. Differential Geom. **33** (1991), 363–378.

[Sa] S. SALAMON - *Riemannian geometry and holonomy groups*, Pitman Research Notes in Math., no. 201, Longman Scientific & Technical, Essex, 1989.

[Scho] J. SCHOUTEN - *On the number of degrees of freedom of the geodetically moving systems*, Proc. Kon. Acad. Wet. Amsterdam **21** (1918), 607–613.

[Schw] L. SCHWACHHÖFER - *On the classification of holonomy representations*, Habilitationsschrift, Universität Leipzig, 1998.

[Wi] B. WILKING - *On compact Riemannian manifolds with noncompact holonomy groups*, 1999 preprint.
URL: http://wwwmath.uni-muenster.de/math/inst/sfb/about/publ/wilking.ps

Robert BRYANT

Duke University
Department of Mathematics
P.O. Box 90320
DURHAM, NC 27708-0320
USA

E–mail : bryant@math.duke.edu

Astérisque

YURI I. MANIN

Classical computing, quantum computing, and Shor's factoring algorithm

Astérisque, tome 266 (2000), Séminaire Bourbaki,
exp. n° 862, p. 375-404

<http://www.numdam.org/item?id=SB_1998-1999__41__375_0>

© Société mathématique de France, 2000, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (http://smf4.emath.fr/Publications/Asterisque/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Séminaire BOURBAKI
51ème année, 1998-99, n° 862

Juin 1999

CLASSICAL COMPUTING, QUANTUM COMPUTING, AND SHOR'S FACTORING ALGORITHM

by **Yuri I. MANIN**

0. WHY QUANTUM COMPUTING ?

Information processing (computing) is the dynamical evolution of a highly organized physical system produced by technology (computer) or nature (brain). The initial state of this system is (determined by) its input; its final state is the output. Physics describes nature in two complementary modes: classical and quantum. Up to the nineties, the basic mathematical models of computing, Turing machines, were classical objects, although the first suggestions for studying quantum models date back at least to 1980.

Roughly speaking, the motivation to study quantum computing comes from several sources: physics and technology, cognitive science, and mathematics. We will briefly discuss them in turn.

(i) Physically, the quantum mode of description is more fundamental than the classical one. In the seventies and eighties it was remarked that, because of the superposition principle, it is computationally unfeasible to simulate quantum processes on classical computers ([Po], [Fe1]). Roughly speaking, quantizing a classical system with N states we obtain a quantum system whose state space is an $(N-1)$–dimensional complex projective space whose volume grows *exponentially* with N. One can argue that the main preoccupation of quantum chemistry is the struggle with resulting difficulties. Reversing this argument, one might expect that quantum computers, if they can be built at all, will be considerably more powerful than classical ones ([Fe1], [Ma2]).

Progress in the microfabrication techniques of modern computers has already led us to the level where quantum noise becomes an essential hindrance to the error–free functioning of microchips. It is only logical to start exploiting the essential quantum

mechanical behavior of small objects in devising computers, instead of neutralizing it.

(ii) As another motivation, one can invoke highly speculative, but intriguing, conjectures that our brain is in fact a quantum computer. For example, the recent progress in writing efficient chess playing software (Deep Blue) shows that to simulate the world championship level using only classical algorithms, one has to be able to analyze about 10^6 positions/sec and use about 10^{10} memory bytes. Since the characteristic time of neuronal processing is about 10^{-3} sec, it is very difficult to explain how the classical brain could possibly do the job and play chess as successfully as Kasparov does. A less spectacular, but not less resource consuming task, is speech generation and perception, which is routinely done by billions of human brains, but still presents a formidable challenge for modern computers using classical algorithms.

Computational complexity of cognitive tasks has several sources: basic variables can be fields; a restricted amount of small blocks can combine into exponentially growing trees of alternatives; databases of incompressible information have to be stored and searched.

Two paradigms have been developed to cope with these difficulties: logic–like languages and combinatorial algorithms, and statistical matching of observed data to an unobserved model (see D. Mumford's paper [Mu] for a lucid discussion of the second paradigm.)

In many cases, the second strategy efficiently supports an acceptable performance, but usually cannot achieve excellency of the Deep Blue level. Both paradigms require huge computational resources, and it is not clear, how they can be organized, unless hardware allows massive parallel computing.

The idea of "quantum parallelism" (see sec. 2 below) is an appealing theoretical alternative. However, it is not at all clear that it can be made compatible with the available experimental evidence, which depicts the central nervous system as a distinctly classical device.

The following way out might be worth exploring. The implementation of efficient quantum algorithms which have been studied so far can be provided by one, or several, quantum chips (registers) controlled by a classical computer. A very considerable part of the overall computing job, besides controlling quantum chips, is also assigned to the classical computer. Analyzing a physical device of such architecture, we would have direct access to its classical component (electrical or neuronal network), whereas locating its quantum components might constitute a considerable challenge. For example, quantum chips in the brain might be represented by macromolecules

of the type that were considered in some theoretical models for high temperature superconductivity.

The difficulties are seemingly increased by the fact that quantum measurements produce non–deterministic outcomes. Actually, one could try to use this to one's advantage, because there exist situations where we can distinguish the quantum randomness from the classical one by analyzing the probability distributions and using the Bell–type inequalities. With hindsight, one recognizes in Bell's setup the first example of the game–like situation where quantum players can behave demonstrably more efficiently that the classical ones (cf. the description of this setup in [Ts], pp. 52–54).

It would be extremely interesting to devise an experimental setting purporting to show that some fragments of the central nervous system relevant for information processing can in fact be in a quantum superposition of classical states.

(iii) Finally, we turn to mathematics. One can argue that nowadays one does not even need additional motivation, given the predominant mood prescribing the quantization of "everything that moves". Quantum groups, quantum cohomology, quantum invariants of knots etc come to mind. This actually seemed to be the primary motivation before 1994, when P. Shor ([Sh]) devised the first quantum algorithm showing that prime factorization can be done on quantum computers in polynomial time, that is, considerably faster than by any known classical algorithm. (P. Shor's work was inspired by the earlier work [Si] of D. Simon). Shor's paper gave a new boost to the subject. Another beautiful result due to L. Grover ([Gro]) is that a quantum search among N objects can be done in $c\sqrt{N}$ steps. A. Kitaev [Ki1] devised new quantum algorithms for computing stabilizers of abelian group actions; his work was preceded by that of D. Boneh and R. Lipton [BoL], who treated the more general problem by a modification of Shor's method (cf. also [Gri]). At least as important as the results themselves, are the tools invented by Shor, Grover, and Kitaev.

Shor's work is the central subject of this lecture. It is explained in sec. 4. This explanation follows the discussion of the general principles of quantum computing and massive quantum parallelism in sec. 2, and of four quantum subroutines, including Grover's searching algorithm, in sec. 3. The second of these subroutines involving quantum computations of classical computable functions shows how to cope with the basic issue of quantum reversibility vs classical irreversibility. For more on this, see [Ben1] and [Ben2]. The opening sec. 1 contains a brief report on the classical theory of computability. I made some effort to express certain notions of computer science, including P/NP, in the language of mainstream mathematics. The last section 5

discusses Kolmogorov complexity in the context of classical and quantum computations.

Last, but not least, the hardware for quantum computing does not exist as yet: see 3.3 below for a brief discussion of the first attempts to engineer it. The quantum algorithms invented and studied up to now will stimulate the search of technological implementation which – if successful – will certainly correct our present understanding of quantum computing and quantum complexity.

Acknowledgements. I am grateful to Alesha Kitaev, David Mumford, and Dimitri Manin for their interest and remarks on the earlier version of this report. Many of their suggestions are incorporated in the text.

1. CLASSICAL THEORY OF COMPUTATION

1.1. Constructive universe. In this section I deal only with deterministic computations, which can be modelled by classical discrete time dynamical systems and subsequently quantized.

Alan Turing undertook the microscopic analysis of the intuitive idea of algorithmic computation. In a sense, he found its genetic code. The atom of information is one bit, the atomary operators can be chosen to act upon one/two bits and to produce the outputs of the same small size. Finally, the sequence of operations is strictly determined by the local environment of bounded size, again several bits.

For a change, I proceed in the reverse direction, and start this section with a presentation of the macrocosm of the classical theory of computation. Categorical language is appropriate to this end.

Let $\mathcal{C}$ be a category whose objects are countable or finite sets U. Elements x of these sets will generally be finite sets with additional structure. Without waiting for all the necessary axioms to be introduced, we will call $x \in U$ *a constructive object of type* U (an integer, a finite graph, a word in a given alphabet, a Boolean expression, an instance of a mass problem ...) The set U itself will be called *the constructive world* of objects of fixed type, and $\mathcal{C}$ *the constructive universe.* The category $\mathcal{C}$, which will be made more concrete below, will contain all finite products and finite unions of its objects, and also finite sets U of all cardinalities.

Morphisms $U \to V$ in $\mathcal{C}$ are certain partial maps of the underlying sets. More precisely, such a morphism is a pair $(D(f), f)$ where $D(f) \subset U$ and $f : D(f) \to V$ is a set–theoretic map. Composition is defined by

$$(D(g), g) \circ (D(f), f) = (g^{-1}D(f), g \circ f).$$

We will omit $D(f)$ when it does not lead to a confusion.

The morphisms f that we will be considering are *(semi)computable functions* $U \to V$. An intuitive meaning of this notion, which has a very strong heuristic potential, can be explained as follows: there should exist an algorithm φ such that if one takes as input the constructive object $u \in U$, one of the three alternatives holds:

(i) $u \in D(f)$, φ produces in a finite number of steps the output $f(u) \in V$.

(ii) $u \notin D(f)$, φ produces in a finite number of steps the standard output meaning NO.

(iii) $u \notin D(f)$, φ works for an infinitely long time without producing any output.

The necessity of including the alternative (iii) in the definition of (semi–)computability was an important and non–trivial discovery of the classical theory. The set of all morphisms $U \to V$ is denoted $\mathcal{C}(U, V)$.

The sets of the form $D(f) \subset U$ are called *enumerable subsets of* U. If both $E \subset U$ and $U \setminus E$ are enumerable, E is called *decidable*.

The classical computation theory makes all of this more precise in the following way.

1.2. Definition. *A category $\mathcal{C}$ as above is called a constructive universe if it contains the constructive world* $\mathbf{N}$ *of all integers* ≥ 1, *finite sets* $\emptyset$, $\{1\}, \ldots, \{1, \ldots, n\}, \ldots$ *and satisfies the following conditions* (a)–(d).

(a) $\mathcal{C}(\mathbf{N}, \mathbf{N})$ *is defined as the set of all partially recursive functions (see e.g.* [Ma1], *Chapter V, or* [Sa]).

(b) *Any infinite object of C is isomorphic to* $\mathbf{N}$.

(c) *If U is finite, $\mathcal{C}(U, V)$ consists of all partial maps $U \to V$. If V is finite, $\mathcal{C}(U, V)$ consists of such f that inverse image of any element of V is enumerable.*

Before stating the last condition (d), we make some comments.

Statement (b) is a part of the famous Church Thesis. Any isomorphism (computable bijection) $\mathbf{N} \to U$ in $\mathcal{C}$ is called a *numbering*. Thus, two different numberings of the same constructive world differ by a recursive permutation of $\mathbf{N}$. We will call such numberings equivalent ones. Notice that because of (c) two finite constructive worlds are isomorphic iff they have the same cardinality, and the automorphism group of any finite U consists of all permutations of U.

As a matter of principle, we always consider $\mathcal{C}$ as an open category, and at any moment allow ourselves to add to it new constructive worlds. If some infinite V is added to $\mathcal{C}$, it must come together with a class of equivalent numberings. Thus, any finite union of constructive worlds can be naturally turned into the constructive

world, so that the embeddings become computable morphisms, and their images are decidable. As another example, the world $\mathbf{N}^*$ of finite sequences of numbers from $\mathbf{N}$ ("words in alphabet $\mathbf{N}$") is endowed with Gödel's numbering

$$(n_1, n_2, \ldots, n_k, \ldots) \mapsto 2^q 3^{n_1 - 1} \ldots p_{k+1}^{n_k - 1} - 1 \tag{1}$$

where p_k is the k–th prime number, $q = \max\{i \mid n_k = \ldots = n_{k-i+1} = 1\}$. Hence we may assume that $\mathcal{C}$ is closed with respect to the construction $U \mapsto U^*$. All natural functions, such as length of the word $U^* \to \mathbf{N}$, or the i–th letter of the word $U^* \to U$ are computable.

Similarly, $\mathcal{C}$ can be made closed with respect to the finite direct products by using the (inverse) numbering of $\mathbf{N}^2$:

$$(m, n) \mapsto m + \frac{1}{2}(m + n - 1)(m + n - 2). \tag{2}$$

Projections, diagonal maps, fiber maps $V \to U \times V, v \mapsto (u_0, v)$ are all computable.

Decidable subsets of constructive worlds are again constructive.

Church Thesis is often invoked as a substitute for an explicit construction of a numbering, and it says that *the category $\mathcal{C}$ is defined uniquely up to equivalence.*

We now turn to the computability properties of the sets of morphisms $\mathcal{C}(U, V)$. Again, it is a matter of principle that $\mathcal{C}(U, V)$ itself *is not a constructive world if U is infinite.* To describe the situation axiomatically, consider first any diagram

$$\mathrm{ev} : P \times U \to V \tag{3}$$

in $\mathcal{C}$. It defines a partial map $P \to \mathcal{C}(U, V)$, $p \mapsto \overline{p}$, where $\overline{p}(u) := \mathrm{ev}\,(p, u)$. We will say that the constructive world $P = P(U, V)$ together with the evaluation map ev is *a programming method* (for computing some maps $U \to V$). It is called *universal,* if the following two conditions are satisfied. First, the map $P \to \mathcal{C}(U, V)$ must be surjective. Second, for any programming method $Q = Q(U, V)$ with the same source U and target V, $\mathcal{C}(Q, P)$ contains translation morphisms

$$\mathrm{trans} : Q(U, V) \to P(U, V) \tag{4}$$

which are, by definition, everywhere defined, computable maps $Q \to P$ such that if $q \mapsto p$, then $\overline{q} = \overline{p}$.

We now complete the Definition 1.2 by adding the last axiom forming part of the Church Thesis:

(d) *For every two constructive worlds U, V, there exist universal programming methods.*

The standard examples of P for $U = V = \mathbf{N}$ are (formalized descriptions of) Turing machines, or recursive functions.

From (d) it follows that the composition of morphisms can be lifted to a computable function on the level of programming methods. To be more precise, if Q (resp. P) is a programming method for U, V (resp. V, W), and R is a universal programming method for U, W, there exist computable composition maps

$$\text{comp} : P(V,W) \times Q(U,V) \to R(U,W),\ (p,q) \mapsto r \tag{5}$$

such that $\overline{r} = \overline{p} \circ \overline{q}$.

Concrete $P(U,V)$ are furnished by the choice of what is called the "model of computations" in computer science. This last notion comes with a detailed description not only of programs but also of all steps of the computational process. At this stage the models of kinematics and dynamics of the process first emerge, and the discussion of quantization can start.

A formalized description of the first n steps will be called *a history of computation* or, for short, *a protocol* (of length n.) For a fixed model, protocols (of all lenghts) form a constructive world as well. We will give two formalized versions of this notion, for functions with infinite and finite domains respectively. The first will be well suited for the discussion of polynomial time computability, the second is the base for quantum computing.

1.3. Models of computations I: normal models. Let U be an infinite constructive world. In this subsection we will be considering partial functions $U \to U$. The more general case $U \to V$ can be reduced to this one by working with $U \coprod V$.

A normal model of computations is the structure $(P, U, I, F, s,)$ consisting of four sets and a map:

$$I \subset U,\ F \subset P \times U,\ s : P \times U \to P \times U\,. \tag{6}$$

Here s is an everywhere defined function such that $s(p,u) = (p, s_p(u))$ for any $(p,u) \in P \times U$. Intuitively, p is a program, u is a configuration of the deterministic discrete time computing device, and $s_p(u)$ is the new configuration obtained from u after one unit of time (clock tick). Two additional subsets $I \subset U$ (initial configurations, or inputs) and $F \subset P \times U$ (final configurations) must be given, such that if $(p,u) \in F$, then $s(p,u) = (p,u)$ i.e., u is a fixed point of s_p.

In this setting, we denote by f_p the partial function $f_p : I \to U$ such that we have

(7) $\quad u \in D(f_p)$ and $f_p(u) = v$ iff for some $n \geq 0$, $(p, s_p^n(u)) \in F$ and $s_p^n(u) = v$.

The minimal such n will be called the *time* (number of clock ticks) needed to calculate $f_p(u)$ using the program p.

Any finite sequence

$$(p, u, s_p(u), \ldots, s_p^m(u)),\ u \in I, \tag{8}$$

will be called *a protocol of computation of length* m.

We now add the constructivity conditions.

We require P, U to be constructive worlds, s computable. In addition, we assume that I, F are decidable subsets of U, $P \times U$ respectively. Then f_p are computable, and protocols of given length (resp. of arbitrary length, resp. or those stopping at F), form constructive worlds. If we denote by Q the world of protocols stopping at F and by $\mathrm{ev} : Q \times U \to U$ the map $(p, u) \mapsto s_p^{\max}(u)$, we get a programming method.

Such a model is called *universal*, if the respective programming method is universal.

The notion of normal model of computations generalizes both normal algorithms and Turing machines. For their common treatment see e.g. [Sa], Chapter 4. In broad terms, $p \in P$ is the list of Markov substitutions, or the table defining the operation of a Turing machine. The remaining worlds U, I, F consist of various words over the working alphabet.

1.3.1. Claim. *For any U, universal normal models of computations exist, and can be effectively constructed.*

For $U = \mathbf{N}$, this follows from the existence of universal Turing machines, and generally, from the Church Thesis. It is well known that the universal machine for calculating functions of k arguments is obtained by taking an appropriate function of $k + 1$ arguments and making the first argument the variable part of the program. Hence P, in this case, consists of pairs (q, m), where q is the fixed program of the $(k + 1)$–variable universal function (hardware) and m is a word written on the tape (software).

1.4. Models of computations II: Boolean circuits. Boolean circuits are classical models of computation well suited for studying maps between the finite sets whose elements are encoded by sequences of 0's and 1's.

Consider the Boolean algebra $\mathbf{B}$ generated over $\mathbf{F}_2$ by a countable sequence of independent variables, say $x_1, x_2, x_3, \ldots$ This is the quotient algebra of $\mathbf{F}_2[x_1, x_2, \ldots]$ with respect to the relations $x_i^2 = x_i$. Each Boolean polynomial determines a function on $\bigoplus_{i=1}^{\infty} \mathbf{F}_2$ with values in $\mathbf{F}_2 = \{0, 1\}$.

We start with the following simple fact.

1.4.1. Claim. *Any map $f : \mathbf{F}_2^m \to \mathbf{F}_2^n$ can be represented by a unique vector of Boolean polynomials.*

Proof. It suffices to consider the case $n = 1$. Then f is represented by

$$F(x_1, \ldots, x_n) := \sum_{y=(y_i)\in\mathbf{F}_2^m} f(y) \prod_i (x_i + y_i + 1) \tag{9}$$

because the product in (9) is the delta function in x supported by y. Moreover, the spaces of maps and of Boolean polynomials have the common dimension 2^m over $\mathbf{F}_2$.

Now we can calculate any vector of Boolean polynomials iterating operations from a small finite list, which is chosen and fixed, e.g. $\mathcal{B} := \{x,\ 1,\ x + y,\ xy,\ (x, x)\}$. Such operators are called *classical gates.* A sequence of such operators, together with indication of their arguments from the previously computed bits, is called *a Boolean circuit.* The number of steps in such a circuit is considered as (a measure of) the time of computation.

When the relevant finite sets are not $\mathbf{F}_2^m$ and perhaps have a wrong cardinality, we encode their elements by finite sequences of bits and consider the restriction of the Boolean polynomial to the relevant subset.

As above, a protocol of computation in this model can be represented as the finite table consisting of rows (generally of variable length) which accommodate sequences of 0's and 1's. The initial line of the table is the input. Each subsequent line must be obtainable from the previous one by the application of one the basic functions in $\mathcal{B}$ to the sequence of neighboring bits (the remaining bits are copied unchanged). The last line is the output. The exact location of the bits which are changed in each row and the nature of change must be a part of the protocol.

Physically, one can implement the rows as the different registers of the memory, or else as the consecutive states of the same register (then we have to make a prescription for how to cope with the variable length, e.g. using blank symbols).

1.4.2. Turing machines vs Boolean circuits. Any protocol of the Turing computation of a function can be treated as such a protocol of an appropriate Boolean circuit, and in this case we have only one register (the initial part of the tape) whose

states are consecutively changed by the head/processor. We will still use the term "gate" in this context.

A computable function f with infinite domain is the limit of a sequence of functions f_i between finite sets whose graphs extend each other. A Turing program for f furnishes a computable sequence of Boolean circuits, which compute all f_i in turn. Such a sequence is sometimes called *uniform.*

1.5. Size, complexity, and polynomial time computability. The quantitative theory of computational models deals simultaneously with the space and time dimensions of protocols. The preceding subsection focused on time, here we introduce space. For Boolean (and Turing machine) protocols this is easy: the length of each row of the protocol is the space required at that moment (plus several more bits for specifying the next gate). The maximum of these lengths is the total space required.

The case of normal models and infinite constructive worlds is more interesting.

Generally we will call *a size function* $U \to \mathbf{N} : u \to |u|$ any function such that for every $B \in \mathbf{N}$, there are only finitely many objects with $|u| \leq B$. Thus the number of bits $|n| = [\log_2 n] + 1$ and the identical function $\|n\| = n$ are both size functions. Using a numbering, we can transfer them to any constructive world. In these two examples, the number of constructive objects of size $\leq H$ grows as $\exp cH$, resp. cH. Such a count in more general cases allows one to make a distinction between *the bit size,* measuring the length of a description of the object, and *the volume* of the object.

In most cases we require computability of size functions. However, there are exceptions: for example, Kolmogorov complexity is a non–computable size function with very important properties: see below and sec. 5.

Given a size function (on all relevant worlds) and a normal model of computations $\mathcal{S}$, we can consider the following complexity problems.

(A) *For a given morphism (computable map)* $f : U \to V$*, estimate the smallest size* $K_{\mathcal{S}}(f)$ *of the program* p *such that* $f = f_p$.

Kolmogorov, Solomonoff and Chaitin proved that there exists an *optimal* universal model of computations $\mathcal{U}$ such that, with $P = \mathbf{N}$ and the bit size function, for any other model $\mathcal{S}$ there exists a constant c such that for any f

$$K_{\mathcal{U}}(f) \leq K_{\mathcal{S}}(f) + c.$$

When $\mathcal{U}$ is chosen, $K_{\mathcal{U}}(f)$ is called Kolmogorov's complexity of f. With a different choice of $\mathcal{U}$ we will get the same complexity function up to $O(1)$–summand.

This complexity measure is highly non–trivial (and especially interesting) for an one–element world U and infinite V. It measures then the size of the most compressed

description of a variable constructive object in V. This complexity is quite "objective" being almost independent of any arbitrary choices. Being uncomputable, it cannot be directly used in computer science. However, it furnishes some basic restrictions on various complexity measures, somewhat similar to those provided by the conservation laws in physics.

On $\mathbf{N}$ we have $K_{\mathcal{U}}(n) \leq |n| + O(1) = \log_2\|n\| + O(1)$. The first inequality "generically" can be replaced by equality, but infinitely often $K_{\mathcal{U}}(n)$ becomes much smaller that $|n|$.

(B) *For a given morphism (recursive map) $f : U \to V$, estimate the time needed to calculate $f(u), u \in D(f)$ using the program p and compare the results for different p and different models of computations.*

(C) *The same for the function "maximal size of intermediate configurations in the protocol of the computation of $f(u)$ using the program p" (space, or memory).*

In the last two problems, we have to compare functions rather than numbers: time and space depend on the size of input. Here a cruder polynomial scale appears naturally. Let us show how this happens.

Fix a computational model $\mathcal{S}$ with the transition function s computing functions $U \to U$, and choose a bit size function on U satisfying the following crucial assumption:

(•) $|u| - c \leq |s_p(u)| \leq |u| + c$ *where the constant c may depend on p but not on u.*

In this case we have $|s_p^m(u)| \leq |u| + c_p m$: the required space grows no more than linearly with time.

Let now $(\mathcal{S}', s')$ be another model such that $s_p = s'_q$ for some q. For example, such q always exists if $\mathcal{S}'$ is universal. Assume that s' satisfies (•) as well, and additionally

(••) *s can be computed in the model $\mathcal{S}'$ in time bounded by a polynomial F in the size of input.*

This requirement is certainly satisfied for Turing and Markov models, and is generally reasonable, because an elementary step of an algorithm deserves its name only if it is computationally tractable.

Then we can replace one application of s_p to $s_p^m(u)$ by $\leq F(|u| + cm)$ applications of s'_q. And if we needed $T(u)$ steps in order to calculate $f_p(u)$ using $\mathcal{S}$, we will need no more than $\leq \sum_{m=1}^{T(u)} F(|u| + cm)$ steps to calculate the same function using $\mathcal{S}'$ and q. In a detailed model, there might be a small additional cost of merging two protocols. This is an example of the translation morphism (4) lifted to the worlds of protocols.

Thus, from (•) and (••) it follows that functions computable in polynomial time by $\mathcal{S}$ have the same property for all reasonable models. Notice also that for such functions, $|f(u)| \leq G(|u|)$ for some polynomial G and that the domain $D(f)$ of such a function is decidable: if after $T(|u|)$ s_p–steps we are not in a final state, then $u \notin D(f)$.

Thus we can define the class PF of functions, say, $\mathbf{N}^k \to \mathbf{N}$ computable in polynomial time by using a fixed universal Turing machine and arguing as above that this definition is model–independent.

If we want to extend it to a constructive universe $\mathcal{C}$ however, we will have to postulate additionally that any constructive world U comes together with a natural class of numberings which, together with their inverses, are computable in polynomial time. This seems to be a part of the content of the *"polynomial Church thesis"* invoked by M. Freedman in [Fr1]. If we take this strengthening of the Church thesis for granted, then we can define also the bit size of an arbitrary constructive object as the bit size of its number with respect to one of these numberings. The quotient of two such size functions is bounded from above and from zero.

Below we will be considering only the universes $\mathcal{C}$ and worlds U with these properties, and $|u|$ will always denote one of the bit size norms. Gödel's numbering (2) for $\mathbf{N} \times \mathbf{N}$ shows that such $\mathcal{C}$ is still closed with respect to finite products. (Notice however that the beautiful numbering (3) of $\mathbf{N}^*$ using primes is not polynomial time computable; it may be replaced by another one which is in PF).

1.6. P/NP **problem.** By definition, a subset $E \subset U$ *belongs to the class* P iff its characteristic function χ_E (equal to 1 on E and 0 outside) belongs to the class PF. Furthermore, $E \in U$ *belongs to the class* NP iff there exists a subset $E' \subset U \times V$ belonging to P and a polynomial G such that

$$u \in E \iff \exists\, (u, v) \in E' \text{ with } |v| \leq G(|u|).$$

Here V is another world (which may coincide with U). We will say that E is obtained from E' by a *polynomially truncated projection.*

The discussion above establishes in what sense this definition is model independent.

Clearly, $P \subset NP$. The inverse inclusion is highly problematic. A naive algorithm calculating χ_E from $\chi_{E'}$ by searching for v with $|v| \leq G(|u|)$ and $\chi_{E'}(u, v) = 1$ will take exponential time e.g. when there is no such v (because $|u|$ is a bit size function). Of course, if one can treat all such v in parallell, the required time will be polynomial. Or else, if an oracle tells you that $u \in E$ and supplies an appropriate

v, you can convince yourself that this is indeed so in polynomial time, by computing $\chi_{E'}(u, v) = 1$.

Notice that the enumerable sets can be alternatively described as projections of decidable ones, and that in this context projection does create undecidable sets. Nobody was able to translate the diagonalization argument used to establish this to the P/NP domain. M. Freedman ([Fr2]) suggested an exciting new approach to the problem $P \neq NP(?)$, based upon a modification of Gromov's strategy for describing groups of polynomial growth.

It has long been known that this problem can be reduced to checking whether some very particular sets – NP–complete ones – belong to P. The set $E \subset U$ is called *NP–complete* if, for any other set $D \subset V, D \in NP$, there exists a function $f : V \to U, f \in PF$, such that $D = f^{-1}(E)$, that is, $\chi_D(v) = \chi_E(f(v))$. We will sketch the classical argument (due to S. Cooke, L. Levin, R. Karp) showing the existence of NP–complete sets. In fact, the reasoning is constructive: it furnishes a polynomially computable map producing f from the descriptions of $\chi_{E'}$ and of the truncating polynomial G.

In order to describe one NP–complete problem, we will define an infinite family of Boolean polynomials b_u indexed by the following data, constituting objects u of the constructive world U. One u is a collection

$$m \in \mathbf{N}; \quad (S_1, T_1), \ldots, (S_N, T_N), \tag{10}$$

where S_i, $T_i \subset \{1, \ldots, m\}$, and b_u is defined as

$$b_u(x_1, \ldots, x_m) = \prod_{i=1}^{N} \left(1 + \prod_{k \in S_i} (1 + x_k) \prod_{j \in T_i} x_j\right). \tag{11}$$

The size of (10) is by definition $|u| = mN$.

Put

$$E = \{u \in U \mid \exists v \in \mathbf{F}_2^m,\ b_u(v) = 1\}.$$

Using the language of Boolean truth values, one says that v *satisfies* b_u if $b_u(v) = 1$, and E is called the *satisfiability problem*, or SAT.

1.6.1. Claim. $E \in NP$.

In fact, let

$$E' = \{(u, v) \mid b_u(v) = 1\} \subset U \times \left(\bigoplus_{i=1}^{\infty} \mathbf{F}_2\right). \tag{12}$$

Clearly, E is the full projection of E'. A contemplation will convince the reader that $E' \in P$. In fact, we can calculate $b_u(v)$ performing $O(Nm)$ Boolean multiplications and additions. The projection to E can be replaced by a polynomially truncated projection, because we have to check only v of size $|v| \leq m$.

1.6.2. Claim. *E is NP–complete.*

In fact, let $D \in NP$, $D \subset A$ where A is some universe. Take a representation of D as a polynomially truncated projection of some set $D' \subset A \times B, D' \in P$. Choose a normal, say Turing, model of computation and consider the Turing protocols of computation of $\chi_{D'}(a, b)$ with fixed a and variable polynomially bounded b. As we have explained above, for a given a, any such protocol can be imagined as a table of a fixed polynomially bounded size whose rows are the consecutive states of the computation. In the "microscopic" description, the positions in this table can be filled only by 0 or 1. In addition, each row is supplied by the specification of the position and the inner state of the head/processor. Some of the arrangements are valid protocols, others are not, but the local nature of the Turing computation allows one to produce a Boolean polynomial b_u in appropriate variables such that the valid protocols are recognized by the fact that this polynomial takes value 1. For detailed explanations see e.g. [GaJ], sec. 2.6. This defines the function f reducing D to E. The construction is so direct that the polynomial time computability of f is straightforward.

Many natural problems are known to be NP–complete, in particular 3–SAT. It is defined as the subset of SAT consisting of those u for which $\operatorname{card}(S_i \cup T_i) = 3$ for all i.

1.6.3. Remark. Most of Boolean functions are not computable in polynomial time. Several versions of this statement can be proved by simple counting.

First of all, fix a finite basis $\mathcal{B}$ of Boolean operations as in 1.4.1, each acting upon $\leq a$ bits. Then sequences of these operations of length t generate $O((bn^a)^t)$ Boolean functions $\mathbf{F}_2^n \to \mathbf{F}_2^n$ where $b = \operatorname{card} \mathcal{B}$. On the other hand, the number of all functions 2^{n2^n} grows as a double exponential of n and for large n cannot be obtained in time t polynomially bounded in n.

The same conclusion holds if we consider not all functions but only permutations: Stirling's formula for $\operatorname{card} S_{2^n} = 2^n!$ involves a double exponential.

Here is one more variation of this problem: define the time complexity of a conjugacy class in S_{2^n} as the minimal number of steps needed to calculate some permutation in this class. This notion arises if we are interested in calculating automorphisms of a finite universe of cardinality 2^n, which is not supplied with a specific encoding by

binary words. Then it can happen that a judicious choice of encoding will drastically simplify the calculation of a given function. However, for most functions we still will not be able to achieve polynomial type computability, because the asymptotical formula for the number of conjugacy classes (partitions)

$$p(2^n) \sim \frac{\exp\left(\pi\sqrt{\frac{2}{3}(2^n - \frac{1}{24})}\right)}{4\sqrt{3}(2^n - \frac{1}{24})}$$

again displays the double exponential growth.

2. QUANTUM PARALLELISM

In this section we will discuss the basics: how to use the superposition principle in order to accelerate (certain) classical computations.

2.1. Description of the problem. *Let N be a large number, $F : \{0, \dots, N-1\} \to \{0, \dots, N-1\}$ a function such that the computation of each particular value $F(x)$ is tractable, that is, can be done in time polynomial in $\log x$. We want to compute (to recognize) some property of the graph $(x, F(x))$, for example:*

(i) *Find the least period r of F, i.e., the least residue $r \bmod N$ such that $F(x + r \bmod N) = F(x)$ for all x (the key step in the Factorization Problem).*

(ii) *Find some x such that $F(x) = 1$ or establish that such x does not exist (Search Problem).*

As we already mentioned, the direct attack on such a problem consists in compiling the complete list of pairs $(x, F(x))$ and then applying to it an algorithm recognizing the property in question. Such a strategy requires at least exponential time (as a function of the bit size of N) since already the length of the list is N. Barring a theoretical breakthrough in understanding such problems (for example a proof that $P = NP$), a practical response might be in exploiting the possibility of parallel computing, i.e., calculating simultaneously many – or even all – values of $F(x)$. This takes less time but uses (dis)proportionally more hardware.

A remarkable suggestion due to D. Deutsch (see [DeuJ], [Deu]) consists in using a quantum superposition of the classical states $|x\rangle$ as the replacement of the union of N classical registers, each in one of the initial states $|x\rangle$. To be more precise, here is a mathematical model formulated as the definition.

2.2. Quantum parallel processing: version I. *Keeping the notation above, assume moreover that $N = 2^n$ and that F is a bijective map (the set of all outputs is a permutation of the set of all inputs).*

(i) *The quantum space of inputs/outputs is the* 2^n*-dimensional complex Hilbert space* H_n *with the orthonormal basis* $|x\rangle$, $0 \leq x \leq N-1$. *Vectors* $|x\rangle$ *are called classical states.*

(ii) *The quantum version of* F *is the unique unitary operator* $U_F : H_n \to H_n$ *such that* $U_F|x\rangle = |F(x)\rangle$.

Quantum parallel computing of F *is (a physical realization of) a system with the state space* H_n *and the evolution operator* U_F.

Naively speaking, if we apply U_F to the initial state which is a superposition of all classical states with, say, equal amplitudes, we will get simultaneously all classical values of F (i.e., their superposition):

$$U_F\left(\frac{1}{\sqrt{N}}\sum|x\rangle\right) = \frac{1}{\sqrt{N}}\sum|F(x)\rangle. \tag{14}$$

We will now discuss various issues related to this definition, before passing to its more realistic modification.

(A) We put $N = 2^n$ above because we are imagining the respective classical system as an n–bit register: cf. the discussion of Boolean circuits. Every number $0 \leq x \leq N-1$ is written in the binary notation $x = \sum_i \epsilon_i 2^i$ and is identified with the pure (classical) state $|\epsilon_{n-1}, \ldots, \epsilon_0\rangle$ where $\epsilon_i = 0$ or 1 is the state of the i–th register. The quantum system H_1 is called *qubit*. We have $H_n = H_1^{\otimes n}$, $|\epsilon_{n-1}, \ldots, \epsilon_0\rangle = |\epsilon_{n-1}\rangle \otimes \ldots \otimes |\epsilon_0\rangle$.

This conforms to the general principles of quantum mechanics. The Hilbert space of the union of systems can be identified with the tensor product of the Hilbert spaces of the subsystems. Accordingly, decomposable vectors correspond to the states of the compound for which one can say that the individual subsystems are in definite states.

(B) Pure quantum states, strictly speaking, are points of the projective space $P(H_n)$ that is, complex lines in H_n. Traditionally, one considers instead vectors of norm one. This leaves undetermined an overall phase factor $\exp i\varphi$. If we have two state vectors, individual phase factors have no objective meaning, but their quotient, that is the difference of their phases, does have one. This difference can be measured by observing effects of interference. This possibility is used for implementing efficient quantum algorithms.

(C) If a quantum system S is isolated, its dynamical evolution is described by the unitary operator $U(t) = \exp iHt$ where H is the Hamiltonian, t is time. Therefore

one option for implementing U_F physically is to design a device for which U_F would be a fixed time evolution operator. However, this seemingly contradicts many deeply rooted notions of the algorithm theory. For example, calculating $F(x)$ for different inputs x takes different times, and it would be highly artificial to try to equalize them already in the design.

Instead, one can try to implement U_F as the result of a sequence of brief interactions, carefully controlled by a classical computer, of S with environment (say, laser pulses). Mathematically speaking, U_F is represented as a product of some standard unitary operators $U_m \dots U_1$ each of which acts only on a small subset (two, three) of classical bits. These operators are called *quantum gates.*

The complexity of the respective quantum computation is determined by its length (the number m of the gates) and by the complexity of each of them. The latter point is a subtle one: continuous parameters, e.g. phase shifts, on which U_i may depend, makes the information content of each U_i potentially infinite and leads to a suspicion that a quantum computer will in fact perform an analog computation, only implemented in a fancy way. A very interesting discussion in [Ts], Lecture 9, convincingly refutes this viewpoint, by displaying those features of quantum computation which distinguish it from both analog and digital classical information processing. This discussion is based on the technique of fault tolerant computing using quantum codes for producing continuous variables highly protected from external noise.

(D) From the classical viewpoint, the requirement that F must be a permutation looks highly restrictive (for instance, in the search problem F takes only two values). Physically, the reason for this requirement is that only such F extend to unitary operators ("quantum reversibility"). The standard way out consists of introducing *two* n–bit registers instead of one, for keeping the value of the argument as well as that of the function. More precisely, if $F(|x\rangle)$ is an arbitrary function, we can replace it by the permutation $\widetilde{F}(|x,y\rangle) := |x, F(x) \oplus y\rangle$, where $\oplus$ is the Boolean (bitwise) sum. This involves no more than a polynomial increase of the classical complexity, and the restriction of $\widetilde{F}$ to $y = 0$ produces the graph of F which we need anyway for the type of problems we are interested in.

In fact, in order to process a classical algorithm (sequence of Boolean gates) for computing F into the quantum one, we replace each classical gate by the respective reversible quantum gate, i.e., by the unitary operator corresponding to it tensored by the identical operator. Besides two registers for keeping $|x\rangle$ and $F(|x\rangle)$ this trick introduces as well extra qubits in which we are not particularly interested. The corresponding space and its content are sometimes referred to as "scratchpad", "garbage",

etc. Besides ensuring reversibility, additional space and garbage can be introduced as well for considering functions $F : \{0, \dots, N-1\} \to \{0, \dots, M-1\}$ where N, M are not powers of two (then we extend them to the closest power of two). For more details, see the next section.

Notice that the choice of gate array (Boolean circuit) as the classical model of computation is essential in the following sense: a quantum routine cannot use conditional instructions. Indeed, to implement such an instruction we must observe the memory in the midst of calculation, but the observation generally will change its current quantum state.

In the same vein, we must avoid copying instructions, because the classical copying operator $|x\rangle \to |x\rangle \otimes |x\rangle$ is not linear. In particular, each output qubit from a quantum gate can be used only in one gate at the next step (if several gates are used parallelly): cloning is not allowed.

These examples show that the basics of quantum code writing will have a very distinct flavor.

We now pass to the problems posed by the input/output routines.

Input, or initialization, in principle can be implemented in the same way as a computation: we produce an input state starting e.g. from the classical state $|0\rangle$ and applying a sequence of basic unitary operators: see the next section. Output, however, involves an additional quantum mechanical notion: that of *observation.*

(E) The simplest model of observation of a quantum system with the Hilbert space H involves the choice of an orthonormal basis of H. Only elements of this basis $|\chi_i\rangle$ can appear as the results of observation. If our system is in some state $|\psi\rangle$ at the moment of observation, it will be observed in the state $|\chi_i\rangle$ with probability $|\langle\chi_i|\,\psi\rangle|^2$.

This means first of all that every quantum computation is inherently probabilistic. Observing (a part of) the quantum memory is not exactly the same as "printing the output". We must plan a series of runs of the same quantum program and the subsequent classical processing of the observed results, and we can hope only to get the desired answer with probability close to one.

Furthermore, this means that by implementing quantum parallelism simplemindedly as in (14), and then observing the memory as if it were the classical n–bit register, we will simply get some value $F(x)$ with probability $1/N$. This does not use the potential of the quantum parallelism. Therefore we formulate a corrected version of this notion, leaving more flexibility and stressing the additional tasks of the designer, each of which eventually contributes to the complexity estimate.

2.3. Quantum parallel processing: version II. *To solve efficiently a problem involving properties of the graph of a function F, we must design:*

(i) *An auxiliary unitary operator U carrying the relevant information about the graph of F.*

(ii) *A computationally feasible realization of U with the help of standard quantum gates.*

(iii) *A computationally feasible realization of the input subroutine.*

(iv) *A computationally feasible classical algorithm processing the results of many runs of quantum computation.*

All of this must be supplemented by quantum error–correcting encoding, which we will not address here. In the next section we will discuss some standard quantum subroutines.

3. SELECTED QUANTUM SUBROUTINES

3.1. Initialization. Using the same conventions as in (14) and the subsequent comments, in particular, the identification $H_n = H_1^{\otimes n}$, we have

$$\frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle = \frac{1}{\sqrt{N}} \sum_{\epsilon_i=0,1} |\epsilon_{n-1} \dots \epsilon_0\rangle = \left(\frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \right)^{\otimes n} . \tag{15}$$

In other words,

$$\frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle = U_1^{(n-1)} \dots U_1^{(0)} |0 \dots 0\rangle \tag{16}$$

where $U_1 : H_1 \to H_1$ is the unitary operator

$$|0\rangle \mapsto \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle), \ |1\rangle \mapsto \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) ,$$

and $U_1^{(i)} = \mathrm{id} \otimes \dots \otimes U_1 \otimes \dots \otimes \mathrm{id}$ acts only on the i–th qubit.

Thus making the quantum gate U_1 act on each memory bit, one can in n steps initialize our register in the state which is the superposition of all 2^n classical states with equal weights.

3.2. Quantum computations of classical functions. Let $\mathcal{B}$ be a finite basis of classical gates containing one–bit identity and generating all Boolean circuits, and

$F : \mathbf{F}_2^m \to \mathbf{F}_2^n$ a function. We will describe how to turn a Boolean circuit of length L calculating F into another Boolean circuit of comparable length consisting only of reversible gates, and calculating a modified function, which however contains all information about the graph of F. Reversibility means that each step is a bijection (actually, an involution) and hence can be extended to a unitary operator, that is, a quantum gate. For a gate f, define $\widetilde{f}(|x,y\rangle) = |x, f(x)+y\rangle$ as in 2.2(D) above.

3.2.1. Claim. *A Boolean circuit $\mathcal{S}$ of length L in the basis $\mathcal{B}$ can be processed into the reversible Boolean circuit $\widetilde{\mathcal{S}}$ of length $O((L+m+n)^2)$ calculating a permutation $H : \mathbf{F}_2^{m+n+L} \to \mathbf{F}_2^{m+n+L}$ with the following property:*

$$H(x,y,0) = (x, F(x)+y, 0) = (\widetilde{F}(x,y), 0).$$

Here x, y, z have sizes m, n, L respectively.

Proof. We will understand L here as the sum of sizes of the outputs of all gates involved in the description of $\mathcal{S}$. We first replace in $\mathcal{S}$ each gate f by its reversible counterpart $\widetilde{f}$. This involves inserting extra bits which we put side by side into a new register of total length L. The resulting subcircuit will calculate a permutation $K : \mathbf{F}_2^{m+L} \to \mathbf{F}_2^{m+L}$ such that $K(x,0) = (F(x), G(x))$ for some function G (garbage).

Now add to the memory one more register of size n keeping the variable y. Extend K to the permutation $\overline{K} : \mathbf{F}_2^{m+L+n} \to \mathbf{F}_2^{m+L+n}$ keeping y intact: $\overline{K} : (x,0,y) \mapsto (F(x), G(x), y)$. Clearly, $\overline{K}$ is calculated by the same Boolean circuit as K, but with extended register.

Extend this circuit by the one adding the contents of the first and the third register: $(F(x), G(x), y) \mapsto (F(x), G(x), F(x)+y)$. Finally, build the last extension which calculates $\bar{K}^{-1}$ and consists of reversed gates calculating $\overline{K}$ in reverse order. This clears the middle register (scratchpad) and produces $(x, 0, F(x)+y)$. The whole circuit requires $O(L+m+n)$ gates if we allow the application of them to not necessarily neighboring bits. Otherwise we must insert gates for local permutations which will replace this estimate by $O((L+m+n)^2)$.

3.3. Fast Fourier transform. Finding the least period of a function of one real variable can be done by calculating its Fourier transforms and looking at its maxima. The same strategy is applied by Shor in his solution of the factorization problem. We will show now that the discrete Fourier transform Φ_n is computationally easy (quantum polynomial time). We define $\Phi_n : H_n \to H_n$ by

$$\Phi_n(|x\rangle) = \frac{1}{\sqrt{N}} \sum_{c=0}^{N-1} |c\rangle \exp(2\pi i c x/N) \tag{17}$$

In fact, it is slightly easier to implement directly the operator

$$\Phi_n^t(|x\rangle) = \frac{1}{\sqrt{N}} \sum_{c=0}^{N-1} |c^t\rangle \exp(2\pi i c x/N). \tag{18}$$

where c^t is c read from the right to the left. The effects of the bit reversal can be then compensated at a later stage without difficulty.

Let $U_2^{(kj)} : H_n \to H_n$, $k < j$, be the quantum gate which acts on the pair of the k–th and j–th qubits in the following way: it multiplies $|11\rangle$ by $\exp(i\pi/2^{j-k})$ and leaves the remaining classical states $|00\rangle, |01\rangle, |10\rangle$ intact.

3.3.1. Lemma *We have*

$$\Phi_n^t = \prod_{k=0}^{n-1} \left(U_1^{(k)} \prod_{j=k+1}^{n-1} U_2^{(kj)} \right). \tag{19}$$

By our rules of the game, (19) has polynomial length in the sense that it involves only $O(n^2)$ gates. However, implementation of $U_2^{(kj)}$ requires controlling variable phase factors which tend to 1 as $k-j$ grows. Moreover, arbitrary pairs of qubits must allow quantum mechanical coupling so that for large n the interaction between qubits must be non–local. The contribution of these complications to the notion of complexity cannot be estimated without going into the details of physical arrangement. Therefore I will add a few words to this effect.

The implementation of quantum register suggested in [CZ] consists of a collection of ions (charged atoms) in a linear harmonic trap (optical cavity). Two of the electronic states of each ion are denoted $|0\rangle$ and $|1\rangle$ and represent a qubit. Laser pulses transmitted to the cavity through the optical fibers and controlled by the classical computer are used to implement gates and read out. The Coulomb repulsion keeps ions apart (spatial selectivity) which allows the preparation of each ion separately in any superposition of $|0\rangle$ and $|1\rangle$ by timing the laser pulse properly and preparing its phase carefully. The same Coulomb repulsion allows for collective excitations of the whole cluster whose quanta are called phonons. Such excitations are produced by laser pulses as well under appropriate resonance conditions. The resulting resonance selectivity combined with the spatial selectivity implements a controlled entanglement of the ions that can be used in order to simulate two and three bit gates. For a detailed and lucid mathematical explanation, see [Ts], Lecture 8.

Another recent suggestion ([GeC]) is to use a single molecule as a quantum register, representing qubits by nuclear spins of individual atoms, and using interactions

through chemical bonds in order to perform multiple bit logic. The classical technique of nuclear magnetic resonance developed since the 1940's, which allows one to work with many molecules simultaneously, provides the start up technology for this project.

3.4. Quantum search. All the subroutines described up to now boiled down to some identities in the unitary groups involving products of not too many operators acting on subspaces of small dimension. They did not involve output subroutines and therefore did not "compute" anything in the traditional sense of the word. We will now describe the beautiful quantum search algorithm due to L. Grover which produces a new identity of this type, but also demonstrates the effect of observation and the way one can use quantum entanglement in order to exploit the potential of quantum parallelism.

We will treat only the simplest version. Let $F : \mathbf{F}_2^n \to \{0,1\}$ be a function taking the value 1 at exactly one point x_0. We want to compute x_0. We assume that F is computable in polynomial time, or else that its values are given by an oracle. Classical search for x_0 requires on the average about $N/2$ evaluations of F where $N = 2^n$.

In the quantum version, we will assume that we have a quantum Boolean circuit (or quantum oracle) calculating the unitary operator $H_n \to H_n$

$$I_F : |x\rangle \mapsto e^{\pi i F(x)}|x\rangle.$$

In other words, I_F is the reflection inverting the sign of $|x_0\rangle$ and leaving the remaining classical states intact.

Moreover, we put $J = -I_\delta$, where $\delta : \mathbf{F}_2^n \to \{0,1\}$ takes the value 1 only at 0, and $V = U_1^{(n-1)} \dots U_1^{(0)}$, as in (16).

3.4.1. Claim. (i) *The real plane in H_n spanned by the uniform superposition ξ of all classical states* (15) *and by $|x_0\rangle$ is invariant with respect to $T := VJVI_F$.*

(ii) *T restricted to this plane is the rotation (from ξ to $|x_0\rangle$) by the angle φ_N where*

$$\cos\varphi_N = 1 - \frac{2}{N}, \ \sin\varphi_N = 2\frac{\sqrt{N-1}}{N}.$$

The check is straightforward.

Now, φ_N is close to $\frac{2}{\sqrt{N}}$, and for the initial angle φ between ξ and $|x_0\rangle$ we have

$$\cos\varphi = -\frac{1}{\sqrt{N}}.$$

Hence in $[\varphi/\varphi_N] \approx \frac{\pi\sqrt{N}}{4}$ applications of T to ξ we will get the state very close to $|x_0\rangle$. Stopping the iteration of T after as many steps and measuring the outcome in the basis of classical states, we will obtain $|x_0\rangle$ with probability very close to one.

One application of T replaces in the quantum search one evaluation of F. Thus, thanks to quantum parallelism, we achieve a polynomial speed–up in comparison with the classical search. The case when F takes value 1 at several points and we only want to find one of them, can be treated by an extension of this method. If there are n such points, the algorithm requires about $\sqrt{N/n}$ steps, and n needs not be known a priori: see [BoyBHT].

4. SHOR'S FACTORING ALGORITHM

4.1. Notation. Let M be a number to be factored. We will assume that it is odd and is not a power of a prime number.

Denote by N the size of the basic memory register we will be using (not counting scratchpad). Its bit size n will be about twice that of M. More precisely, choose $M^2 < N = 2^n < 2M^2$. Finally, let $1 < t < M$ be a random parameter with $\gcd(t, M) = 1$. This condition can be checked classically in time polynomial in n.

Below we will describe one run of Shor's algorithm, in which t (and of course, M, N) is fixed. Generally, polynomially many runs will be required, in which the value of t can remain the same or be chosen anew. This is needed in order to gather statistics. Shor's algorithm is a probabilistic one, with two sources of randomness that must be clearly distinguished. One is built into the classical probabilistic reduction of factoring to the finding of the period of a function. Another stems from the necessity of observing quantum memory, which, too, produces random results.

More precise estimates than those given here show that a quantum computer which can store about $3n$ qubits can find a factor of M in time of order n^3 with probability close to 1 : see [BCDP]. On the other hand, it is widely believed that no recursive function of the type $M \mapsto$ *a proper factor of* M belongs to PF. This is why the most popular public key encryption schemes rely upon the difficulty of the factoring problem.

4.2. Classical algorithm. Put

$$r := \min\{\rho \mid t^\rho \equiv 1 \bmod M\}$$

which is the least period of $F : a \mapsto t^a \bmod M$.

4.2.1. Claim. *If one can efficiently calculate r as a function of t, one can find a proper divisor of M in polynomial in $\log_2 M$ time with probability $\geq 1 - M^{-m}$ for any fixed m.*

Assume that for a given t the period r satisfies

$$r \equiv 0 \bmod 2, \ t^{r/2} \not\equiv -1 \bmod M.$$

Then $\gcd(t^{r/2}+1, M)$ is a proper divisor of M. Notice that gcd is computable in polynomial time.

The probability that this condition holds is $\geq 1 - \frac{1}{2^{k-1}}$ where k is the number of different odd prime divisors of M, hence $\geq \frac{1}{2}$ in our case. Therefore we will find a good t with probability $\geq 1 - M^{-m}$ in $O(\log M)$ tries. The longest calculation in one try is that of $t^{r/2}$. The usual squaring method takes polynomial time as well.

4.3. Quantum algorithm calculating r. Here we describe one run of the quantum algorithm which purports to compute r, given M, N, t. We will use the working register that can keep a pair consisting of a variable $0 \leq a \leq N-1$ and the respective value of the function $t^a \bmod M$. One more register will serve as the scratchpad needed to compute $|a, t^a \bmod M\rangle$ reversibly. When this calculation is completed, the content of the scratchpad will be reversibly erased: cf. 3.2.1. In the remaining part of the computation the scratchpad will not be used anymore, we can decouple it, and forget about it.

The quantum computation consists of four steps, three of which were described in sec. 3:

(i) Partial initialization produces from $|0,0\rangle$ the superposition

$$\frac{1}{\sqrt{N}} \sum_{a=0}^{N-1} |a, 0\rangle.$$

(ii) Reversible calculation of F processes this state into

$$\frac{1}{\sqrt{N}} \sum_{a=0}^{N-1} |a, t^a \bmod M\rangle.$$

(iii) Partial Fourier transform then furnishes

$$\frac{1}{N} \sum_{a=0}^{N-1} \sum_{c=0}^{N-1} \exp(2\pi i a c/N) \, |c, t^a \bmod M\rangle.$$

(iv) The last step is the observation of this state with respect to the system of classical states $|c, m \bmod M\rangle$. This step produces some concrete output

$$|c, t^k \bmod M\rangle \tag{20}$$

with probability

$$\left| \frac{1}{N} \sum_{a:\, t^a \equiv t^k \bmod M} \exp(2\pi i a c/N) \right|^2 . \tag{21}$$

The remaining part of the run is assigned to the classical computer and consists of the following steps.

(A) *Find the best approximation (in lowest terms) to* $\frac{c}{N}$ *with denominator* $r' < M < \sqrt{N}$*:*

$$\left| \frac{c}{N} - \frac{d'}{r'} \right| < \frac{1}{2N}. \tag{22}$$

As we will see below, we may hope that r' will coincide with r in at least one run among at most polynomially many. Hence we try r' in the role of r right away:

(B) *If* $r' \equiv 0 \bmod 2$, *calculate* $\gcd(t^{r'/2} \pm 1, M)$.

If r' is odd, or if r' is even, but we did not get a proper divisor of M, repeat the run $O(\log \log M)$ times with the same t. In case of failure, change t and start a new run.

4.3.1. Justification. We will now show that, given t, from the observed values of $|c, t^k \bmod M\rangle$ in $O(\log \log M)$ runs we can find the correct value of r with probability close to 1.

Let us call the observed value of c *good*, if

$$\exists l \in \left[-\frac{r}{2}, \frac{r}{2}\right], \; rc \equiv l \bmod N.$$

In this case there exists such d that

$$-\frac{r}{2} \leq rc - dN = l \leq \frac{r}{2}$$

so that

$$\left| \frac{c}{N} - \frac{d}{r} \right| < \frac{1}{2N}.$$

Hence if c is good, then r' found from (22) in fact divides r.

Now call c *very good* if $r' = r$.

Estimating the exponential sum (21), we can easily check that the probability of observing a good c is $\geq \frac{1}{3r^2}$. On the other hand, there are $r\varphi(r)$ states $|c, t^k \bmod M\rangle$ with very good c. Thus to find a very good c with high probability, $O(r^2 \log r)$ runs will suffice.

5. KOLMOGOROV COMPLEXITY AND GROWTH OF RECURSIVE FUNCTIONS

Consider general functions $f : \mathbf{N} \to \mathbf{N}$. Computability theory uses several growth scales for such functions, of which two are most useful: f may be majorized by some recursive function (e.g. when it is itself recursive), or by a polynomial (e.g. when it is computable in polynomial time). Linear growth does not seem particularly relevant in this context. However, this impression is quite misleading, at least if one allows re–ordering $\mathbf{N}$. In fact, we have:

5.1. Claim. *There exists a permutation* $\mathbf{K} : \mathbf{N} \to \mathbf{N}$ *such that for any partially recursive function* $f : \mathbf{N} \to \mathbf{N}$ *there exists a constant* c *with the property*

$$\mathbf{K} \circ f \circ \mathbf{K}^{-1}(n) \leq cn \text{ for all } n \in \mathbf{K}(D(f)). \tag{23}$$

Moreover, $\mathbf{K}$ *is bounded by a linear function, but* $\mathbf{K}^{-1}$ *is not bounded by any recursive function.*

Proof. We will use the Kolmogorov complexity measure. For a recursive function $u : \mathbf{N} \to \mathbf{N}$, $x \in \mathbf{N}$, put $C_u(x) := \min\{\mathrm{k} \,|\, \mathrm{f(k)} = \mathrm{x}\}$, or ∞ if such k does not exist. Call such a function u *optimal* if, for any other recursive function v, there exists a constant $c_{u,v}$ such that $C_u(x) \leq c_{u,v} C_v(x)$ for all x. Optimal functions do exist (see e.g. [Ma1], Theorem VI.9.2); in particular, they take all positive integer values (however they certainly are not everywhere defined). Fix one such u and call $C_u(x)$ the (exponential) complexity of x. By definition, $\mathbf{K} = \mathbf{K}_u$ rearranges $\mathbf{N}$ in the order of increasing complexity. In other words,

$$\mathbf{K}(x) := 1 + \mathrm{card}\,\{\mathrm{y} \,|\, \mathrm{C_u(y)} < \mathrm{C_u(x)}\}. \tag{24}$$

We first show that

$$\mathbf{K}(x) = \exp(O(1))\, C_u(x). \tag{25}$$

Since C_u takes each value at most once, it follows from (24) that $\mathbf{K}(n) \leq C_u(n)$. In order to show that $C_u(x) \leq c\,\mathbf{K}(x)$ for some c it suffices to check that

$$\operatorname{card}\{k \leq N \mid \exists\, x,\, C_u(x) = k\} \geq b\,N$$

with some $b > 0$. In fact, at least half of the numbers $x \leq N$ have the complexity which is no less than $x/2$.

Now, VI.9.7(b) in [Ma1] implies that, for any recursive function f and all $x \in D(f)$, we have $C_u(f(x)) \leq \operatorname{const} C_u(x)$. Since $C_u(x)$ and $\mathbf{K}(x)$ have the same order of growth up to a bounded factor, our claim follows.

5.2. Corollary. *Denote by S_∞^{rec} the group of recursive permutations of* $\mathbf{N}$. *Then $\mathbf{K}\, S_\infty^{\mathrm{rec}}\, \mathbf{K}^{-1}$ is a subgroup of permutations of no more than linear growth.*

Actually, appealing to the Proposition VI.9.6 of [Ma1], one can considerably strengthen this result. For example, let σ be a recursive permutation, $\sigma^{\mathbf{K}} = \mathbf{K}\sigma\mathbf{K}^{-1}$. Then $\sigma^{\mathbf{K}}(x) \leq cx$ so that $(\sigma^{\mathbf{K}})^n(x) \leq c^n x$ for $n > 0$. But actually the last inequality can be replaced by

$$(\sigma^{\mathbf{K}})^n(x) \leq c'n$$

for a fixed x and variable n. With both x and n variable one gets the estimate $O(xn \log(xn))$.

In the same way as finite permutations appear in the quantum versions of Boolean circuits, infinite (computable) permutations are natural for treating quantum Turing machines ([Deu]) and our normal computation models. In fact, if one assumes that the transition function s is a permutation, and then extends it to the unitary operator U_s in the infinite–dimensional Hilbert space, one might be interested in studying the spectral properties of such operators. But the latter depend only on the conjugacy class. Perhaps the universal conjugation $U_{\mathbf{K}}$ might be a useful theoretical tool in this context. In the purely classical situation, (23) may play a role in studying the limiting behavior of polynomial time algorithms, as suggested in [Fr1] and [Fr2].

Finally, I would like to comment upon the hidden role of Kolmogorov complexity in the real life of classical computing. The point is that in a sense (which is difficult to formalize), we are interested only in the calculation of sufficiently nice functions, because a random Boolean function will have (super)exponential complexity anyway. A nice function, at the very least, has a short description and, therefore, a small Kolmogorov complexity. Thus, dealing with practical problems, we actually work not with small numbers, graphs, circuits, ... , but rather with an initial segment of the

respective constructive world reordered with the help of **K**. We systematically replace a large object by its short description, and then try to overcome the computational difficulties generated by this replacement.

APPENDIX

The following text is a contribution to the prehistory of quantum computing. It is the translation from Russian of the last three paragraphs of the Introduction to [Ma2] (1980). For this reference I am grateful to A. Kitaev [Ki].

"Perhaps, for better understanding of this phenomenon [DNA replication], we need a mathematical theory of quantum automata. Such a theory would provide us with mathematical models of deterministic processes with quite unusual properties. One reason for this is that the quantum state space has far greater capacity than the classical one: for a classical system with N states, its quantum version allowing superposition accommodates c^N states. When we join two classical systems, their number of states N_1 and N_2 are multiplied, and in the quantum case we get exponential growth $c^{N_1 N_2}$.

These crude estimates show that the quantum behavior of the system might be much more complex than its classical simulation. In particular, since there is no unique decomposition of a quantum system into its constituent parts, a state of the quantum automaton can be considered in many ways as a state of various virtual classical automata. Cf. the following instructive comment at the end of the article [Po]: 'The quantum–mechanical computation of one molecule of methane requires 10^{42} grid points. Assuming that at each point we have to perform only 10 elementary operations, and that the computation is performed at the extremely low temperature $T = 3.10^{-3} K$, we would still have to use all the energy produced on Earth during the last century.'

The first difficulty we must overcome is the choice of the correct balance between the mathematical and the physical principles. The quantum automaton has to be an abstract one: its mathematical model must appeal only to the general principles of quantum physics, without prescribing a physical implementation. Then the model of evolution is the unitary rotation in a finite dimensional Hilbert space, and the decomposition of the system into its virtual parts corresponds to the tensor product decomposition of the state space. Somewhere in this picture we must accommodate interaction, which is described by density matrices and probabilities."

BIBLIOGRAPHY

[BCDP] D. Beckman, A. N. Chari, Sr. Devabhaktuni, J. Preskill - *Efficient networks for quantum computing.* Phys. Rev. A, 54:2 (1996), 1034–1063.

[Ben1] P. Benioff - *The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines*, J. Stat. Phys., 22 (1980), 563–591.

[Ben2] P. Benioff - *Quantum mechanical Hamiltonian models of Turing machines that dissipate no energy*, Phys. Rev. Lett., 48 (1980), 1581–1585.

[BoL] D. Boneh, R. Lipton - *Quantum cryptoanalysis of hidden linear functions*, Proc. of Advances in Cryptology — CRYPTO '95, Springer LN in Computer Science, vol. 963 (1995), 424–437.

[BoyBHT] M. Boyer, G. Brassard, P. Høyer, A. Tapp - *Tight bounds on quantum searching*, Preprint, 1996.

[CZ] J. Cirac, P. Zoller - *Quantum computation with cold trapped ions*, Phys. Rev. Lett., 74:20 (1995), 4091–4094.

[Deu] D. Deutsch - *Quantum theory, the Church–Turing principle and the universal quantum computer*, Proc. R. Soc. Lond. A 400 (1985), 97–117.

[DeuJ] D. Deutsch, R. Jozsa - *Rapid solutions of problems by quantum computation*, Proc. Roy. Soc. London, Ser. A, 449 (1992), 553–558.

[Fe1] R. Feynman - *Simulating physics with computers*, Int. J. of Theor. Phys., 21 (1982), 467–488.

[Fe2] R. Feynman. *Quantum mechanical computers*, Found. Phys., 16 (1986), 507–531.

[Fr1] M. Freedman - *Topological views on computational complexity*, In Proc. ICM Berlin 1998, vol. II, 453–464.

[Fr2] M. Freedman - *Limit, logic, and computation*, Proc. Nat. Ac. Sci. USA, 95 (1998), 95–97.

[Fr3] M. Freedman - *P/NP, and the quantum field computer*, Proc. Nat. Ac. Sci. USA, 95 (1998), 98–101.

[GaJ] M. Garey, D. Johnson - *Computers and Intractability: A Guide to the Theory of NP–Completeness*, W. H. Freeman and Co., San–Francisco, 1979.

[GeC] N. Gershenfield, I. Chuang - *Bulk spin–resonance quantum computation*, Science 275 (1997), 350–355.

[Gri] D. Grigoriev - *Testing the shift–equivalence of polynomials using quantum mechanics*, In: Manin's Festschrift, Journ. of Math. Sci., 82:1 (1996), 3184–3193.

[Gro] L. K. Grover - *Quantum mechanics helps in searching for a needle in a haystack*, Phys. Rev. Lett. 79 (1997), 325–328.

[Ki1] A. Kitaev - *Quantum computations: algorithms and error correction*, Russian Math. Surveys, 52:6 (1997), 53–112.

[Ki2] A. Kitaev - *Classical and quantum computations*, Lecture notes, Independent University, Moscow, 1998.

[Ma1] Yu. Manin - *A Course in Mathematical Logic*, Springer Verlag, 1977, pp. xiii+286.

[Ma2] Yu. Manin - *Computable and uncomputable (in Russian)*, Moscow, Sovetskoye Radio, 1980.

[Mu] D. Mumford - *The statistical description of visual signals*, Preprint.

[Po] R. P. Poplavskii - *Thermodynamical models of information processing (in Russian)*, Uspekhi Fizicheskikh Nauk, 115:3 (1975), 465–501.

[Sa] A. Salomaa - *Computation and Automata*, Cambridge UP, 1985.

[Sh] P. W. Shor - *Polynomial–time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM J. Comput., 26:5 (1997), 1484–1509.

[Si] D. Simon - *On the power of quantum computation*, Proc. of the 35th Ann. Symp. on Foundations of Comp. Sci. (1994), 116–123.

[Ts] B. Tsirelson - *Quantum information processing*, Lecture notes, Tel–Aviv University, 1997.

Yu. I. MANIN

Max–Planck–Institut für Mathematik
Vivatgasse 7
D-53111 BONN (Germany)
E-mail : manin@mpim-bonn.mpg.de

Astérisque

JOSEPH OESTERLÉ

Densité maximale des empilements de sphères en dimension 3

Astérisque, tome 266 (2000), Séminaire Bourbaki,
exp. n° 863, p. 405-413

<http://www.numdam.org/item?id=SB_1998-1999__41__405_0>

© Société mathématique de France, 2000, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (http://smf4.emath.fr/Publications/Asterisque/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Séminaire BOURBAKI Juin 1999
51ème année, 1998-99, n° 863

DENSITÉ MAXIMALE DES EMPILEMENTS DE SPHÈRES EN DIMENSION 3 [d'après Thomas C. HALES et Samuel P. FERGUSON]

par **Joseph OESTERLÉ** (*)

Au nouvel an 1610, Johannes Kepler offre à son protecteur et ami Johannes Matthaüs Wackher von Wackenfels un opuscule d'une vingtaine de pages ([14]), intitulé *Strena seu de nive sexangula* (l'étrenne ou la neige sexangulaire). Il y médite sur l'origine de la forme des flocons de neige en étoile à six branches, ce qui l'amène à s'intéresser entre autres à la forme des alvéoles des abeilles, au dodécaèdre rhomboédrique, à la forme des grains de grenades et à l'empilement de sphères que nous appelons aujourd'hui *cubique à faces centrées*, et qui est par exemple celui suivant lequel les artilleurs empilent les boulets de canon. Kepler affirme qu'il n'est pas possible de trouver un empilement de sphères (toutes de même rayon) plus serré, assertion qui est devenue la fameuse *conjecture de Kepler.*

En 1997, T. Hales publie une stratégie pour démontrer la conjecture de Kepler, en la ramenant à un problème d'optimisation non linéaire en dimension finie. Il mène ce programme à bien, en collaboration avec son étudiant S. Ferguson, et l'achève en juillet 1998. Tant le choix du problème d'optimisation considéré que sa résolution s'appuient sur un usage intensif de l'ordinateur. Mais près de 4 siècles après avoir été formulée, la conjecture de Kepler est enfin démontrée :

THÉORÈME 1 (Hales – Ferguson). — *Tout empilement de sphères dans l'espace euclidien de dimension* 3 *a une densité au plus égale à la densité* $\frac{\pi}{3\sqrt{2}}$ *de l'empilement cubique à faces centrées.*

Précisons le vocabulaire utilisé : Le mot *sphère* signifie ici sphère pleine, *i.e.* est synonyme de boule fermée. Un *empilement de sphères* est une famille de boules fermées, toutes de même rayon, d'intérieurs mutuellement disjoints. Soit B la réunion

(*) L'auteur remercie J. C. Lagarias de lui avoir communiqué ses notes personnelles sur le sujet.

de ces boules ; la *densité de l'empilement* est la limite supérieure de $\mathrm{vol}(B \cap \Omega)/\mathrm{vol}(\Omega)$, pour Ω un cube dont le côté tend vers l'infini.

On peut donner du th. 1 la formulation équivalente suivante, qui ne fait intervenir que des empilements finis :

THÉORÈME 1′. — *Une famille de boules disjointes, toutes de même rayon et contenues dans un parallélépipède Ω de l'espace euclidien de dimension 3, occupe une fraction du volume de Ω au plus égale à $\frac{\pi}{3\sqrt{2}}$.*

Le fait que le th. 1′ implique le th. 1 résulte de la définition de la densité d'un empilement de sphères ; l'implication réciproque se déduit de ce que l'on peut paver l'espace par des translatés d'un parallélépipède.

L'empilement cubique à faces centrées est ainsi nommé car on l'obtient en plaçant les centres des sphères aux sommets et aux centres des faces des mailles d'un réseau cubique (le rayon des sphères étant choisi de façon à ce qu'elles soient tangentes à leurs plus proches voisines). En prenant pour verticale la direction soit d'un côté, soit d'une diagonale de la maille cubique, on peut voir cet empilement comme une superposition de couches horizontales de sphères réparties soit en réseau carré (chaque sphère étant tangente à 4 autres de la même couche), soit en réseau hexagonal (chaque sphère étant tangente à 6 autres de la même couche) ; chaque couche s'insère dans les trous les plus profonds de la précédente.

Remarque 1. — Lorsqu'on superpose des couches de sphères réparties comme ci-dessus en réseau hexagonal, chaque couche peut être posée de deux façons distinctes dans les trous les plus profonds de la précédente. Cela conduit à une infinité (non dénombrable) d'empilements de sphères différents, tous de même densité $\frac{\pi}{3\sqrt{2}}$ que l'empilement cubique à faces centrées.

1. Quelques mots sur l'histoire de la conjecture de Kepler

En 1831, dans son étude des formes quadratiques ternaires, C. F. Gauss ([6]) démontre le cas particulier suivant de la conjecture de Kepler : si les centres des sphères d'un empilement forment un réseau, la densité de l'empilement est $\leq \frac{\pi}{3\sqrt{2}}$, avec égalité si et seulement si l'empilement est cubique à faces centrées. Ce résultat sera largement généralisé en dimension supérieure (cf. [1], [15]).

La conjecture de Kepler est reprise et généralisée par Hilbert dans son $18^{\text{ème}}$ problème.

En 1910, A. Thue ([18]) démontre l'analogue en dimension 2 de la conjecture de Kepler : *tout empilement de cercles dans le plan euclidien a une densité au plus égale*

à la densité $\frac{\pi}{2\sqrt{3}}$ *de l'empilement hexagonal.*

En 1953, L. Fejes Tóth ([3]) propose un programme pour démontrer la conjecture de Kepler en la ramenant à un problème d'optimisation non linéaire en dimension finie. Ce problème, toujours ouvert, semble trop complexe pour être résolu avec les moyens de calcul actuels. Mais Fejes Tóth est le premier à deviner l'importance que le développement rapide des ordinateurs aura pour ce type de stratégie.

En 1958, C. A. Rogers ([16]) démontre que la densité supérieure d'un empilement de sphères dans un espace euclidien de dimension n est au plus égale à σ_n, portion du volume d'un simplexe régulier de côté 2 recouverte par les sphères unité centrées en ses sommets. Pour $n = 2$, cela fournit une nouvelle démonstration du théorème de Thue, car on peut paver le plan euclidien par des triangles équilatéraux.

L'espace euclidien de dimension 3 ne peut être pavé par des tétraèdres réguliers. La borne de Rogers, $\sigma_3 = 0,7796...$, est de ce fait moins bonne que celle prédite par la conjecture de Kepler, $\frac{\pi}{3\sqrt{2}} = 0,7404....$ La phrase célèbre de Rogers : *Many mathematicians believe, and all physicists know, that the density cannot exceed* $\frac{\pi}{3\sqrt{2}}$, exprime mieux que toute autre la redoutable subtilité de la conjecture de Kepler, malgré son apparente simplicité.

La borne de Rogers pour la densité des empilements de sphères en dimension 3 sera améliorée par divers auteurs ; avant les travaux de Hales et Ferguson, la meilleure borne connue était 0,7731 (obtenue par D. J. Muder en 1993). Hales mentionne que les travaux de Muder ont influencé les siens.

Dans un article publié en 1993, Wu-yi Hsiang ([13]) affirme démontrer la conjecture de Kepler. Les spécialistes s'accordent à considérer que les preuves de nombreuses assertions n'y sont pas présentées avec suffisamment de détails pour être qualifiées de complètes. De ce fait, l'article de Hsiang est au mieux vu comme un plan d'attaque de la conjecture de Kepler. Je partage cette opinion. Les textes ultérieurs de Hsiang qui m'ont été communiqués, en particulier le manuscrit d'un livre en préparation, ne me semblent pas pour l'instant de nature à modifier ce point de vue.

2. La conjecture du dodécaèdre et le problème des 13 sphères

Considérons un empilement de sphères de rayon 1 dans un espace euclidien E. Soit S l'ensemble des centres des sphères. La *cellule de Voronoi* de s, notée $V(s)$ se compose des points de E qui sont plus proches de s que des autres points de S. C'est un ensemble convexe fermé, localement polyédrique. Les cellules de Voronoi des points de S pavent l'espace E : autrement dit, elles recouvrent E et leurs intérieurs sont mutuellement disjoints. Chacune de ces cellules contient une unique sphère de

l'empilement ; on en déduit facilement que si leur volume est minoré par v, la densité de l'empilement est majorée par b/v, où b est le volume de la boule unité.

Cette observation conduit naturellement au problème suivant : *quel est le volume minimal des cellules de Voronoi des empilements de sphères de rayon* 1 *?*

En dimension 2, le problème a été résolu par Fejes Tóth en 1943 ([2]) : l'aire d'une cellule de Voronoi d'un empilement de cercles de rayon 1 est $\geq 2\sqrt{3}$, avec égalité si et seulement si la cellule est un hexagone régulier circonscrit à un cercle de rayon 1. Cet énoncé implique le théorème de Thue relatif à la densité des empilements de cercles.

En dimension 3, le problème a été résolu par Hales et Mc Laughlin en novembre 1998, par des techniques analogues à celles utilisées pour prouver la conjecture de Kepler. Ils démontrent ainsi la *conjecture du dodécaèdre*, formulée par Fejes Tóth en 1943 (et démontrée par lui dans le cas particulier où les cellules de Voronoi n'ont que 12 faces au plus) :

THÉORÈME 2 (Hales – Mc Laughlin). — *Le volume d'une cellule de Voronoi d'un empilement de sphères de rayon* 1 *dans l'espace euclidien de dimension* 3 *est* $\geq 10\sqrt{130 - 58\sqrt{5}}$, *avec égalité si et seulement si la cellule est un dodécaèdre régulier circonscrit à une sphère de rayon* 1.

Notons que ce théorème implique que la densité d'un empilement de sphères en dimension 3 est majorée par 0,7547, une borne nettement meilleure que celles obtenues avant que ne soit démontrée la conjecture de Kepler.

Remarque 2. — Pour démontrer le théorème 2, on peut supposer que l'empilement considéré est saturé (*i.e.* maximal). Dans ce cas, pour tout $s \in S$, la cellule de Voronoi $V(s)$ est contenue dans la boule de centre s et de rayon 2. Sa forme, et donc son volume, ne dépendent que des points de S dont la distance à s est ≤ 4, et le nombre de ces points est majoré par une constante absolue. La conjecture du dodécaèdre est donc un problème d'optimisation non linéaire : il s'agit de trouver le maximum de la fonction volume sur un espace de configurations qui est une partie compacte d'un espace vectoriel de dimension finie. Une des difficultés vient du fait que cette dimension est de l'ordre de 150 : en effet, on sait seulement que le nombre de faces de $V(s)$ est ≤ 49 et on connait des exemples où il est 44.

En 1694, une controverse oppose Isaac Newton et l'abbé Gregory. Newton pense qu'on ne peut placer qu'au plus 12 sphères identiques (ne s'interpénétrant pas) au contact d'une sphère centrale de même rayon, Gregory pense qu'il est possible d'en placer 13. Cette alternative, le *problème des 13 sphères*, ne sera tranchée qu'en 1953,

par Schütte et Van der Waerden ([17]) : ils donnent raison à Newton. Ils exhibent par ailleurs un arrangement de 13 sphères de rayon 1 tangentes à une sphère centrale de rayon $r = 1,0911....$ On pense que ce rayon est le plus petit possible, mais cela n'a pour l'instant été démontré. Peut-être les techniques de Hales et Ferguson permettront-elles de le faire.

3. Principe d'attaque de la conjecture de Kepler

Pour majorer la densité des empilements de sphères dans un espace euclidien E de dimension 3, il suffit de considérer les empilements de sphères de rayon 1 qui sont saturés (*i.e.* maximaux). Choisissons donc un tel empilement. Notons B la réunion des boules qui le composent, et S l'ensemble de leurs centres : les distances mutuelles des points de S sont ≥ 2, et tout point de E est à une distance < 2 de S.

Si A est une partie de E, notons χ_A sa fonction caractéristique. Supposons que, pour des nombres réels a, b, c tels que $a > \frac{3c}{4\pi}$, on puisse écrire

$$a\chi_B - b\chi_E = \sum_{s\in S} \phi_s \qquad \text{(presque partout)}, \tag{1}$$

où ϕ_s est une fonction intégrable dont la norme L^1 est majorée par une constante indépendante de s, dont le support est contenu dans une boule de centre s et de rayon indépendant de s, et dont l'intégrale satisfait la relation

$$\int \phi_s \leq c. \tag{2}$$

Démontrons que la densité δ de l'empilement satisfait alors l'inégalité

$$(a - \frac{3c}{4\pi})\delta \leq b \tag{3}$$

En intégrant (1) sur un cube Ω de côté r, on obtient lorsque r tend vers l'infini

$$a\mathrm{vol}(B\cap\Omega) - b\mathrm{vol}(\Omega) \leq c\,\mathrm{Card}(S\cap\Omega) + O(r^2) \leq \frac{3c}{4\pi}\mathrm{vol}(B\cap\Omega) + O(r^2).$$

d'où $(a - \frac{3c}{4\pi})\frac{\mathrm{vol}(B\cap\Omega)}{\mathrm{vol}(\Omega)} \leq b + O(r^{-1})$; l'inégalité (3) s'en déduit par passage à la limite supérieure.

Exemples. — 1) La relation (1) est satisfaite lorsque l'on prend $a = 0$, $b = -1$ et $\phi_s = -\chi_{V(s)}$, où $V(s)$ est la cellule de Voronoi de s ; cette cellule est contenue dans la boule de centre s et de rayon 2, puisque l'empilement est supposé saturé. La relation

(2) est satisfaite en prenant $c = -v$, où v est un minorant du volume des cellules de Voronoi. La relation (3) donne l'inégalité $\delta \leq \frac{4\pi}{3v}$ que nous avions rencontrée au n° 2.

2) Soit r un nombre réel > 2. Disons que deux points distincts de S sont voisins si leur distance est $\leq r$. Fejes Tóth propose pour démontrer la conjecture de Kepler de prendre le plus grand r qui garantisse que tout point de S a au plus 12 voisins (probablement 2,0534...), et de poser $\phi_s = -\frac{12-m}{12}\chi_{\mathrm{V}(s)} - \frac{1}{12}\sum \chi_{\mathrm{V}(t)}$, où t parcourt les voisins de s et m est leur nombre. La relation (1) est dans ce cas satisfaite avec $a = 0$ et $b = -1$, mais la détermination de la constante c optimale dans (2) apparaît totalement hors d'atteinte.

3) On appelle *décomposition de Delaunay* associée à l'empilement une décomposition de l'espace E en tétraèdres, dont les sommets appartiennent à S, et dont les boules circonscrites n'ont pas de point intérieur dans S. Il existe au moins une telle décomposition ; celle-ci est unique lorsqu'aucun point de E n'appartient à plus de 4 cellules de Voronoi (ce qui est le cas générique), et est dans ce cas duale de celle de Voronoi. Fixons une décomposition de Delaunay associée à l'empilement et appelons *étoile* de s la réunion $\mathrm{D}(s)$ des simplexes de la décomposition contenant s. Posons $\phi_s = \chi_{\mathrm{B}\cap\mathrm{D}(s)} - \delta_{oct}\chi_{\mathrm{D}(s)}$, où δ_{oct} est la portion du volume d'un octaèdre régulier de côté 2 recouverte par les boules unités centrées en ses sommets. La première tentative de Hales pour démontrer en 1992 la conjecture de Kepler revenait à utiliser ces fonctions ϕ_s, avec $a = 4$ et $b = 4\delta_{oct}$. Mais il s'aperçut que pour certaines étoiles (les prismes pentagonaux), l'intégrale de ϕ_s était trop grande pour conclure.

4. Articulation de la démonstration de Hales et Ferguson

La démonstration de Hales et Ferguson de la conjecture de Kepler se compose de 8 articles, totalisant environ 300 pages, dont l'ordre logique est le suivant :

An overview of the Kepler conjecture, de T.C. Hales ([7]).
A formulation of the Kepler conjecture, de S. Ferguson et T.C. Hales ([5]).
Sphere packings I, de T.C. Hales ([8]).
Sphere packings II, de T.C. Hales ([9]).
Sphere packings III, de T.C. Hales ([10]).
Sphere packings IV, de T.C. Hales ([11]).
Sphere packings V, de S. Ferguson ([4]).
The Kepler conjecture (Sphere packings VI), de T.C. Hales ([12]).

Les premiers articles rédigés furent *Sphere packings I* et *Sphere packings II*. Ils ont été publiés en 1997. Les autres articles sont pour l'instant des prépublications

électroniques, disponibles sur la page personnelle de T. Hales.

Dans *Sphere packings I*, Hales décrit sa stratégie pour démontrer la conjecture de Kepler. Le problème d'optimisation qu'il considère est de même nature que celui considéré dans l'exemple 3 du n° 3, mais s'appuie sur une décomposition de l'espace hybride entre la décomposition de Voronoi et celle de Delaunay. À l'étoile $\mathrm{D}(s)$ d'un point s pour cette décomposition, Hales associe un nombre réel qu'il appelle *le score* de $\mathrm{D}(s)$ et qui est l'intégrale de la fonction ϕ_s dans l'approche du n° 3.

L'ensemble $\mathcal{D}$ des étoiles possibles, pour $s = 0$, a une structure naturelle d'espace compact et le score est une fonction continue sur $\mathcal{D}$. Les étoiles qui proviennent de l'empilement cubique faces centrée et de ses variantes décrites dans la remarque 1 forment deux orbites $\mathcal{D}_0$ et $\mathcal{D}'_0$ sous l'action du groupe orthogonal. Pour démontrer la conjecture de Kepler, il suffit de démontrer que le maximum du score est atteint sur $\mathcal{D}_0 \cup \mathcal{D}'_0$. Pour des raisons de commodité, Hales note $8pt$ la valeur prise par le score sur cet ensemble.

Chaque étoile possède une structure combinatoire, décrite par une carte sphérique. Dans *Sphere packings II*, Hales démontre une propriété forte de maximum local : les étoiles ayant même structure combinatoire que celles de $\mathcal{D}_0 \cup \mathcal{D}'_0$ ont un score $\leq 8pt$, avec égalité si et seulement si elles appartiennent à $\mathcal{D}_0 \cup \mathcal{D}'_0$.

La stratégie consiste alors à démontrer que pour les autres structures combinatoires, le score est $< 8pt$. Les difficultés techniques rencontrées, en particulier dans l'étude des prismes pentagonaux, conduisent Ferguson et Hales à modifier le problème d'optimisation initial, en proposant une nouvelle décomposition de l'espace et une nouvelle fonction score, plus compliquées apparemment mais mieux adaptées à l'étude numérique envisagée. C'est ce qu'ils font dans *A formulation of the Kepler conjecture*, en prenant soin de vérifier que les résultats de *Sphere packings I* et *Sphere packings II* restent valables dans ce nouveau cadre.

Dans *Sphere packings III*, Hales démontre que le score (nouvelle version) de D est $< 8pt$ si la carte associée n'a que des faces triangulaires ou quadrilatérales, et ne provient pas d'un prisme pentagonal (ni de $\mathcal{D}_0$ et $\mathcal{D}'_0$) ; le cas des prismes pentagonaux est traité par Ferguson dans *Sphere packings V*, celui des cartes possédant des faces à 5 côtés ou plus par Hales dans *Sphere packings IV* et *Sphere packings VI*.

Au total, ce sont environ 5000 cartes, recensées par ordinateur, qui interviennent dans le problème. Des méthodes générales de programmation linéaire permettent de se ramener à n'examiner qu'une centaine de cas ; ceux-ci sont traités un à un. Beaucoup des problèmes d'optimisation non linéaires rencontrés sont remplacés par des problèmes d'optimisation linéaires qui les dominent, que l'on sait résoudre par

les méthodes de programmation linéaire. Au total, ce sont près de 100000 problèmes linéaires, en 100 à 200 variables, avec 1000 à 2000 contraintes, qu'il a fallu traiter pour achever la démonstration. Les calculs ont été effectués en utilisant une *arithmétique des intervalles* : chaque nombre réel r est représenté par un couple (a, b) de nombres rationnels qui l'encadrent, de manière à garantir que les erreurs d'arrondis cumulées n'entachent pas la validité des résultats obtenus.

5. La conjecture de Kepler est-elle démontrée ?

La situation n'est pas sans rappeler celle du théorème des 4 couleurs, dont la démonstration par K. Appel et W. Haken en 1977 avait nécessité l'examen de quelques milliers de configurations par ordinateur.

Je ne prétends pas avoir vérifié l'intégralité de la démonstration de Hales et Ferguson, et encore moins les programmes informatiques utilisés. Je ne pense pas que quiconque à part les auteurs l'ait fait.

Ma conviction personnelle que les travaux de Hales et Ferguson fournissent effectivement une preuve de la conjecture de Kepler repose sur les arguments suivants :

— Ni erreurs substantielles, ni points obscurs n'ont été relevés jusqu'à présent par les spécialistes ayant examiné ces travaux.

— Dans toutes les parties du travail que j'ai pris la peine de lire en détail, j'ai trouvé des démonstrations claires et complètes.

— Le recours à l'intuition géométrique, dangereux dans ce type de problème, est réduit au minimum et remplacé par des inégalités précises.

— Les algorithmes employés sont décrits avec soin, et les codes rédigés sont disponibles pour vérification. L'arithmétique des intervalles tient compte des erreurs commises dans l'approximation des nombres réels.

— Si d'aventure, on s'apercevait ultérieurement qu'un cas à examiner a échappé à la sagacité des auteurs, il est vraisemblable que leurs techniques permettraient de le traiter sans devoir introduire d'idées nouvelles.

BIBLIOGRAPHIE

[1] J. H. CONWAY, N. J. A. SLOANE – *Sphere packings, lattices and groups*, Grundlehren der mathematischen Wissenschaften **290**, Springer Verlag, 1988.

[2] L. FEJES TÓTH – *Über die dichteste Kugellagerung*, Math. Z. **48** (1943), 676-684.

[3] L. FEJES TÓTH – *Lagerungen in der Ebene, auf der Kugel und im Raum*, Springer, Berlin, 1953.

[4] S. P. FERGUSON – *Sphere packings V*, thèse, Université du Michigan, 1997.

[5] S. P. FERGUSON, T. C. HALES – *A formulation of the Kepler conjecture*, prépublication.

[6] C. F. GAUSS – *Untersuchungen über die Eigenschaften der positiven ternären quadratischen Formen von Ludwig August Seber*, Göttingische gelehrte Anzeigen, 1831, Juli 9 (= Werke, vol 2, Königlische Gesellschaft der Wissenschaften, Göttingen, 1876, 188-196).

[7] T. C. HALES – *An overview of the Kepler conjecture*, prépublication.

[8] T. C. HALES – *Sphere packings I*, Disc. Comp. Geom. **17** (1997), 1-51.

[9] T. C. HALES – *Sphere packings II*, Disc. Comp. Geom. **18** (1997), 135-149.

[10] T. C. HALES – *Sphere packings III*, prépublication.

[11] T. C. HALES – *Sphere packings IV*, prépublication.

[12] T. C. HALES – *The Kepler conjecture (Sphere packings VI)*, prépublication.

[13] W-Y. HSIANG – *On the sphere packing problem and the proof of Kepler's conjecture*, Intern. J. Math. **93** (1993), 739-831.

[14] J. KEPLER – *Strena seu de nive sexangula*, Frankfurt, Tampach, 1611, 24 pages. (Traduction allemande de F. Rossmann, *Neujahrsgabe*, Berlin, 1943 ; traduction anglaise de C. Hardie, *The six-cornered snowflake*, Oxford, 1966 ; traduction française de R. Halleux, *L'étrenne ou la neige sexangulaire*, librairie J. Vrin et éditions du CNRS, Paris, 1975.)

[15] J. OESTERLÉ – *Empilements de sphères*, Séminaire Bourbaki 1989-90, exposé n° 727, Astérisque **189-190** (1990), 375-397.

[16] C. A. ROGERS – *The packing of equal spheres*, Proc. London Math. Soc. **8** (1958), 609-620.

[17] K. SCHÜTTE, B. L. VAN DER WAERDEN – *Das Problem der dreizehn Kugeln*, Proc. London Math. Ann. **125** (1953), 325-334.

[18] A. THUE – *Über die dichteste Zusammenstellung von kongruenten Kreisen in einer Ebene*, Norske Vid. Selsk. Skr.**1** (1910), 1-9.

Joseph OESTERLÉ

Institut Henri Poincaré
11, rue Pierre et Marie Curie
F-75005 PARIS
Adresse électronique : oesterle@ihp.jussieu.fr

Astérisque

JEAN-PIERRE SERRE

Sous-groupes finis des groupes de Lie

Astérisque, tome 266 (2000), Séminaire Bourbaki,
exp. n° 864, p. 415-430

<http://www.numdam.org/item?id=SB_1998-1999__41__415_0>

© Société mathématique de France, 2000, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (http://smf4.emath.fr/Publications/Asterisque/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Séminaire BOURBAKI
51ème année, 1998-99, n° 864

Juin 1999

SOUS-GROUPES FINIS DES GROUPES DE LIE

par **Jean–Pierre SERRE**

INTRODUCTION

Les sous-groupes finis du groupe des rotations $\mathrm{SO}_3(\mathbf{R})$ sont bien connus. Ce sont :

les groupes cycliques C_n d'ordre $n = 1, 2, \ldots$;

les groupes diédraux D_n d'ordre $2n$, $n = 2, 3, \ldots$;

le groupe alterné Alt_4 d'ordre 12 ;

le groupe symétrique Sym_4 d'ordre 24 ;

le groupe alterné Alt_5 d'ordre 60.

On aimerait avoir une liste analogue pour d'autres groupes de Lie compacts, ou d'autres groupes algébriques (en caractéristique zéro, et même en caractéristique > 0). Ce serait utile pour beaucoup de questions (représentations ℓ-adiques, par exemple). Bien sûr, c'est trop demander, vu que tout groupe fini se plonge dans un groupe unitaire convenable ! On va voir que l'on peut tout de même dire pas mal de choses si l'on se borne à des groupes finis qui sont, soit abéliens, soit simples.

HYPOTHÈSES ET NOTATIONS

Plutôt que de travailler dans la catégorie des groupes de Lie compacts, on préfère se placer dans celle des groupes réductifs complexes. Cela ne change rien : on sait que, si K est un groupe de Lie compact, il possède un complexifié G qui est un groupe réductif sur $\mathbf{C}$; le groupe K est un sous-groupe compact maximal de $G(\mathbf{C})$. Tout sous-groupe fini de $G(\mathbf{C})$ est conjugué à un sous-groupe de K ; de plus, K "contrôle la fusion de K dans $G(\mathbf{C})$" au sens suivant : si A, B sont deux sous-groupes de K, et si $g \in G(\mathbf{C})$ est tel que $gAg^{-1} = B$, il existe un élément g_0 de K tel que $g_0 a g_0^{-1} = gag^{-1}$ pour tout $a \in A$ (cela se déduit de la décomposition de Cartan de $G(\mathbf{C})$).

(Dans le cas particulier $K = \mathrm{SO}_3(\mathbf{R})$, on a $G = \mathrm{PGL}_2$, de sorte que les groupes $C_n, D_n, \ldots, \mathrm{Alt}_5$ s'interprètent comme des sous-groupes finis de $\mathrm{PGL}_2(\mathbf{C})$, c'est-à-dire

comme des groupes finis d'automorphismes de la droite projective.)

Dans ce qui suit, on adoptera le point de vue des groupes algébriques (qui a, entre autres avantages, celui de permettre des réductions modulo p). On fixe un corps k algébriquement clos de caractéristique zéro, ainsi qu'un groupe réductif connexe G défini sur k ; on se permet d'identifier G à $G(k)$. Le cas le plus intéressant est celui où G est "presque simple", i.e. semi-simple à système de racines irréductible ; le groupe adjoint G^{ad} est alors un groupe simple, au sens usuel du terme.

1. LE CAS (PRESQUE) ABÉLIEN

Lorsque $G = \mathrm{PGL}_2$ les sous-groupes abéliens finis de G sont les groupes cycliques C_n et le groupe diédral D_2 qui est abélien élémentaire de type $(2,2)$. Les C_n sont contenus dans un tore maximal, alors que D_2 ne l'est pas ; le nombre premier $p = 2$ joue donc un rôle particulier pour PGL_2. Nous allons trouver une situation analogue dans le cas général.

1.1. Sous-groupes toraux

Un sous-groupe fini A de G est dit *toral* s'il est contenu dans un tore maximal T de G. La structure d'un tel sous-groupe est évidente : si $r = \dim T$ est le rang de G, A peut être engendré par r éléments ; inversement, tout groupe abélien ayant cette propriété est isomorphe à un sous-groupe toral de G.

Soit $N = N_G(T)$ le normalisateur de T dans G. Le quotient $W = N/T$ est le *groupe de Weyl* de G (plus correctement : du couple (G,T)). Ce groupe opère sur T par conjugaison, et il contrôle la fusion de T dans G :

1.1.1. *Si A et B sont des sous-groupes de T, et si $g \in G$ est tel que $gAg^{-1} = B$, il existe $w \in W$ tel que $w(a) = gag^{-1}$ pour tout $a \in A$.*

Cet énoncé est l'exact analogue d'un théorème de Burnside sur les sous-groupes du centre d'un p-groupe de Sylow. Il se démontre de la même manière : on remarque que T et $g^{-1}Tg$ sont des tores maximaux du centralisateur $Z_G(A)$ de A, donc sont conjugués par $Z_G(A)$. Cela permet de remplacer g par un élément de N ; d'où le résultat cherché.

Les groupes abéliens ayant très peu de générateurs sont toraux :

1.1.2. *Soit A un sous-groupe abélien fini de G. Alors A est toral dans chacun des deux cas suivants :*

a) *A est cyclique ;*

b) *G est simplement connexe, et A est engendré par deux éléments.*

Le cas a) est immédiat : tout élément d'ordre fini est semi-simple, donc contenu dans un tore maximal. Dans le cas b), supposons A engendré par x, y. Du fait que G est simplement connexe, le centralisateur $Z_G(x)$ est connexe. Le même argument que dans a) montre qu'il existe un tore maximal T de $Z_G(x)$ qui contient y. Ce tore est un tore maximal de G et il contient x, donc A.

1.2. Plongements dans N

À défaut de pouvoir plonger un groupe abélien fini dans un tore maximal, on peut essayer de le plonger dans le normalisateur d'un tel tore. C'est toujours possible. Plus généralement (cf. Borel-Serre [6], Borel-Mostow [5] et Springer-Steinberg [33], II.5.6) :

1.2.1. *Soit A un sous-groupe fini hyper-résoluble de G. Il existe un tore maximal T de G dont le normalisateur N contient A.*

Rappelons qu'un groupe A est dit hyper-résoluble ("supersolvable") s'il admet une suite de composition :

$$1 = A_0 \subset A_1 \subset \ldots \subset A_n = A,$$

où les A_i sont normaux dans A, et A_i/A_{i-1} est cyclique pour tout $i \geq 1$. On a les implications :

$$\text{abélien} \Longrightarrow \text{nilpotent} \Longrightarrow \text{hyper-résoluble} \Longrightarrow \text{résoluble}.$$

Voici une application simple de 1.2.1 :

1.2.2. *Soit p un nombre premier ne divisant pas l'ordre du groupe de Weyl W. Si A est un p-groupe contenu dans G, A est abélien et toral.*

En effet, on peut supposer, d'après 1.2.1, que A est contenu dans N. Vu l'hypothèse faite sur p, son image dans $W = N/T$ est triviale. Il est donc contenu dans T.

Remarque : Le groupe N est une extension, en général non triviale, de W par T. On trouvera dans Tits ([35],[36]) une description de cette extension, en termes d'un certain groupe fini $N_{\mathbf{Z}}$ défini explicitement par générateurs et relations ; voir aussi Bourbaki, LIE IX, p. 115, exerc. 12.

1.3. Nombres premiers de torsion

(Références : Borel [4], Steinberg [34] et Bourbaki, LIE IX, p. 120-121, exerc. 7 à 12.)

Un nombre premier p est dit *de torsion* (pour G) s'il vérifie les conditions équivalentes suivantes :

a) *Il existe un p-sous-groupe abélien de G qui n'est pas toral.*

a') *Il existe un p-sous-groupe abélien élémentaire de G, de rang ≤ 3, qui n'est pas toral.*

On note $\mathrm{Tors}(G)$ l'ensemble de ces nombres premiers ; d'après 1.2.2, c'est un sous-ensemble de l'ensemble des diviseurs premiers de l'ordre de W. Dans le cas particulier où $G = \mathrm{PGL}_2$, on a $\mathrm{Tors}(G) = \{2\}$.

Le terme de "torsion" provient du résultat suivant, dans lequel je suppose que $k = \mathbf{C}$ (sinon il faut faire intervenir la cohomologie étale) :

1.3.1. (cf. [4],[34]) *Pour que p appartienne à $\mathrm{Tors}(G)$, il faut et il suffit que l'un des groupes d'homologie $H_i(G, \mathbf{Z})$ contienne un élément d'ordre p.*

(Noter qu'il revient au même de considérer l'homologie de $G = G(\mathbf{C})$ ou celle d'un compact maximal K, car $G(\mathbf{C})$ et K ont même type d'homotopie.)

On trouvera dans [4] et [34] une longue liste de propriétés caractérisant les éléments de $\mathrm{Tors}(G)$. En voici quelques-unes :

1.3.2. *On a $\mathrm{Tors}(G) = \mathrm{Tors}(G')$, où G' est le groupe dérivé de G.*

Comme G' est semi-simple, cela ramène l'étude de $\mathrm{Tors}(G)$ au cas où G est semi-simple. Dans ce cas, notons $\overline{G}$ le revêtement universel de G, notons $\pi_1(G)$ le noyau de $\overline{G} \to G$ et soit $\mathrm{Tors}(\pi_1(G))$ l'ensemble des nombres premiers qui divisent l'ordre du groupe fini $\pi_1(G)$. Alors :

1.3.3. *On a $\mathrm{Tors}(G) = \cup_H \mathrm{Tors}(\pi_1(H'))$, où H parcourt les sous-groupes réductifs connexes de G ayant même rang que G.*

1.3.4. *On a $\mathrm{Tors}(G) = \mathrm{Tors}(\overline{G}) \cup \mathrm{Tors}(\pi_1(G))$.*

Cet énoncé ramène la détermination de $\mathrm{Tors}(G)$ au cas où G est simplement connexe. En utilisant 1.3.3, on en déduit (cf. [4],[34]) :

1.3.5. *Supposons G simplement connexe et presque simple. Soit (α_i) une base de son système de racines, soit β la plus grande racine, et écrivons la racine duale $\beta^\vee$ de β sous la forme :*

$$\beta^\vee = \sum n_i \alpha_i^\vee,$$

où les n_i sont des entiers > 0. Alors, pour que p soit de torsion pour G, il faut et il suffit qu'on ait $p \leq \sup(n_i)$.

D'où :

1.3.6. *Supposons G simplement connexe et presque simple. Alors :*

$\mathrm{Tors}(G) = \emptyset$ *si G est de type A_n ou C_n ;*

$\mathrm{Tors}(G) = \{2\}$ *si G est de type B_n $(n \geq 3)$, D_n $(n \geq 4)$ ou G_2 ;*

$\mathrm{Tors}(G) = \{2, 3\}$ *si G est de type F_4, E_6 ou E_7 ;*

$\mathrm{Tors}(G) = \{2, 3, 5\}$ *si G est de type E_8.*

1.4. Exemples de groupes abéliens élémentaires non toraux

(Références : Adams [1], Borel [4], Borel–Serre [6], Cohen-Seitz [10], Steinberg [34] et (surtout) Griess [17].)

Je me borne à deux exemples, l'un relatif à $p = 2$ et l'autre à $p = 5$.

1.4.1. Supposons que -1 appartienne au groupe de Weyl W ; c'est le cas pour les groupes de type A_1, B_n, C_n, D_n (n pair), G_2, F_4, E_7, E_8. Soit $g \in N$ un représentant de l'élément -1 de W. On peut montrer que g^2 est d'ordre 1 ou 2, et appartient au centre de G. Supposons que $g^2 = 1$ (c'est le cas si G est de type adjoint). Soit A le groupe engendré par g et par les éléments d'ordre 2 de T ; c'est un groupe abélien élémentaire d'ordre 2^{n+1}, où n est le rang de G, i.e. la dimension de T. *Ce groupe n'est pas toral* ; on peut même montrer que son centralisateur $Z_G(A)$ est fini.

Lorsque G est PGL_2, le groupe A est le groupe diédral D_2. Lorsque G est de type G_2, F_4 ou E_8, A est d'ordre 2^3, 2^5, 2^9 ; de tels sous-groupes jouent un grand rôle dans la cohomologie (usuelle – ou galoisienne) du groupe G. Noter que, dans ces trois cas, A est un sous-groupe élémentaire *maximal* de G : de façon générale, si G est simplement connexe, les p-sous-groupes abéliens de G sont de rang $\leq n + 1$ si $p = 2$, et de rang $\leq n$ si $p > 2$, cf. Borel [4] et Cohen-Seitz [10].

1.4.2. Le groupe $G = E_8$ contient un élément z d'ordre 5 dont le centralisateur $Z_G(z)$ est de la forme $G_1 \cdot G_2$, où G_1 et G_2 sont isomorphes à SL_5, commutent, et ont pour intersection $\langle z \rangle$ (cela se déduit du diagramme de Dynkin complété de E_8 en remarquant que, si l'on en retranche la racine simple qui a le coefficient 5 dans la plus grande racine, on trouve deux diagrammes de type A_4). Dans $G_1 = \mathrm{SL}_5$, il est facile de trouver des éléments x_1 , y_1 d'ordre 5 tels que $x_1 y_1 x_1^{-1} y_1^{-1} = z$; de même, il existe dans G_2 des éléments x_2 , y_2 d'ordre 5 tels que $x_2 y_2 x_2^{-1} y_2^{-1} = z^{-1}$. Si l'on pose $x = x_1 x_2$ et $y = y_1 y_2$, on constate que le groupe $A = \langle x, y, z \rangle$ est abélien élémentaire d'ordre 5^3. *Ce groupe n'est pas toral* ; on peut même montrer que $Z_G(A)$ est égal à A.

1.5. Relations entre cohomologie galoisienne et sous-groupes non toraux

Qu'il existe de telles relations est connu depuis longtemps. Voici deux exemples :

1.5.1 (Grothendieck [23]). *Les deux propriétés suivantes sont équivalentes* :

a) $\mathrm{Tors}(G) = \emptyset$ (cela équivaut à dire que tout sous-groupe abélien de G est toral).

b) $H^1(K, G) = 0$ *pour toute extension* K *de* k.

(Pour la définition de $H^1(K, G)$, voir par exemple [30].)

Lorsque G est semi-simple, ces propriétés sont satisfaites si et seulement si G est un produit de groupes simplement connexes de type A ou C, cf. § 1.3. Un tel groupe

est parfois dit "spécial".

1.5.2. *Supposons que G soit égal à PGL_2, ou soit de type G_2. Soit A le 2-sous-groupe élémentaire non toral de G défini dans* 1.4.1. *Alors, pour toute extension K de k, l'application $H^1(K,A) \to H^1(K,G)$ est surjective.*

Noter que $H^1(K,A)$ n'est autre que $\mathrm{Hom}(\mathrm{Gal}(\overline{K}/K), A)$. Dans le cas de PGL_2, les éléments de $H^1(K,A)$ peuvent donc s'interpréter comme des couples (λ,μ) d'éléments de K^*/K^{*2} et l'élément correspondant de $H^1(K,\mathrm{PGL}_2)$ est l'algèbre de quaternions définie par deux générateurs i et j soumis aux relations

$$i^2 = \lambda\ , \quad j^2 = \mu\ , \quad ij = -ji.$$

Même chose pour G_2, les quaternions étant remplacés par les octonions.

L'énoncé 1.5.2, pour agréable qu'il soit, ne donne pas de moyen de prouver la non trivialité des éléments de $H^1(K,G)$ ainsi obtenus ; il faut le compléter par la construction d'*invariants cohomologiques*, cf. [30], §§ 6,7 et n et Youssin [29] ont obtenu un résultat bien plus satisfaisant. Pour le formuler, il faut d'abord définir la *dimension essentielle* $\mathrm{ed}(x)$ d'un élément x de $H^1(K,G)$: c'est la borne inférieure des degrés de transcendance sur k des sous-extensions K' de K telles que x appartienne à l'image de $H^1(K',G) \to H^1(K,G)$. (En termes plus géométriques – et plus vagues – c'est le nombre minimum de paramètres dont on a besoin pour écrire le G-torseur x.) La borne supérieure des $\mathrm{ed}(x)$, quand K et x varient, est la *dimension essentielle* de G ; elle est notée $\mathrm{ed}(G)$. Nous pouvons maintenant énoncer le théorème principal de [29] :

1.5.3. *Si G contient un p-sous-groupe abélien élémentaire A dont le centralisateur est fini, on a* $\mathrm{ed}(G) \geq \mathrm{rang}(A)$.

En combinant cet énoncé avec 1.4.1, on obtient :

1.5.4. *On a* $\mathrm{ed}(E_7^{\mathrm{ad}}) \geq 8$ et $\mathrm{ed}(E_8) \geq 9$.

Ainsi, il existe des E_8-torseurs dont la construction exige au moins 9 paramètres !

Remarques

1) L'hypothèse faite sur A dans 1.5.3 est équivalente à dire que A n'appartient à aucun sous-groupe parabolique propre de G. Elle entraîne que A n'est pas toral.

2) L'énoncé démontré dans [29] est plus précis que 1.5.3 ; c'est :

$$\mathrm{ed}(G;p) \geq \mathrm{rang}(A),$$

où $\mathrm{ed}(G;p)$ est la dimension essentielle de G "en p" (i.e. en considérant comme négligeables les extensions de corps de degré premier à p).

3) Les démonstrations de [29] utilisent la *résolution des singularités* (sous forme équivariante). Elles ne s'étendent pas, pour l'instant, aux corps de caractéristique $\neq 0$.

2. LE CAS (PRESQUE) SIMPLE

On va maintenant s'intéresser aux plongements d'un groupe simple (fini, non abélien) dans G.

Il est commode de considérer, plus généralement, les plongements des groupes $\overline{S}$ qui sont des extensions centrales de S (on peut imposer à $\overline{S}$ d'être égal à son groupe dérivé, cela ne change rien). Exemple typique :

$$S = L_2(q) = \mathrm{PSL}_2(\mathbf{F}_q) \quad \text{et} \quad \overline{S} = 2 \cdot L_2(q) = \mathrm{SL}_2(\mathbf{F}_q)\ , \qquad q \text{ impair.}$$

(Les notations $L_2(q)$ et $2 \cdot L_2(q)$ sont celles de l'ATLAS [15].)

Un plongement d'un tel groupe $\overline{S}$ dans G est appelé un *plongement projectif* de S. Le principal avantage de cette notion est la propriété d'invariance suivante : si $G' \to G$ est une isogénie, S a un plongement projectif dans G si et seulement si il a un plongement projectif dans G'.

Lorsque S et G sont donnés, et que G est un groupe classique, l'examen de la table des caractères de S (et de ses extensions centrales) permet de décider si S a un plongement (ou un plongement projectif) dans G ; c'est clair lorsque G est de type A_n, et c'est facile pour les types B_n, C_n, D_n. Une méthode analogue s'applique à G_2 (en utilisant sa représentation irréductible de degré 7 et la forme trilinéaire alternée correspondante), cf. Aschbacher [2] et Cohen–Wales [11]. Les types F_4, E_6, E_7, E_8 sont plus difficiles ; ce n'est que récemment (Griess–Ryba [22]) que la liste des S possibles a été complétée. Avant de donner cette liste (que l'on trouvera au § 2.4), je vais parler du cas $S = L_2(q)$, qui est le seul où l'on ait des énoncés généraux, i.e. valables aussi bien pour les groupes classiques que pour les groupes exceptionnels.

2.1. Plongements projectifs de $L_2(q)$ dans G ; énoncé du résultat

On suppose $q > 2$. Le groupe $S = L_2(q) = \mathrm{PSL}_2(\mathbf{F}_q)$ est alors un groupe simple (sauf si $q = 3$, où c'est le groupe Alt_4), et toute extension centrale $\overline{S}$ de S, égale à son groupe dérivé, est isomorphe, soit à $2 \cdot S = \mathrm{SL}_2(\mathbf{F}_q)$, soit à S (sauf si $q = 4$ ou 9). On va donc s'intéresser aux homomorphismes

$$f : \mathrm{SL}_2(\mathbf{F}_q) \longrightarrow G$$

de noyau égal à 1 ou à (± 1). Un tel homomorphisme sera dit *non dégénéré*.

Écrivons q sous la forme p^e, avec p premier, $e \geq 1$. Le p-groupe de Sylow $U = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$ de $\mathrm{SL}_2(\mathbf{F}_q)$ est isomorphe à $\mathbf{F}_q$. Le groupe $A = f(U)$ est un p-groupe abélien élémentaire de G de rang e. Nous dirons que f est *de type toral* si A est toral au sens du § 1.1. C'est le cas si $e = 1$, ou si $e = 2$ et G est simplement connexe (1.1.2), ou si p n'est pas un nombre premier de torsion pour G.

Nous allons donner un *critère pour l'existence d'un f non dégénéré de type toral.* Supposons G presque simple, de rang r ; soient k_i $(i = 1, \ldots, r)$ les exposants de son groupe de Weyl et soient $d_i = k_i + 1$ les degrés correspondants (Bourbaki, LIE V, § 6, prop. 3). L'énoncé suivant résume une série de résultats dus à divers auteurs ([2], [8], [9], [11], [12], [14], [19], [20], [24], [28], [31]) :

2.1.1. *Pour qu'il existe un homomorphisme non dégénéré de type toral de* $\mathrm{SL}_2(\mathbf{F}_q)$ *dans G, il faut et il suffit que $q - 1$ divise l'un des entiers $2d_1, \ldots, 2d_r$.*

Remarque.— Lorsque $p = 2$ ou 3, il existe quelques plongements de $L_2(p^e)$ qui ne sont pas de type toral, par exemple :

$$L_2(4) \longrightarrow \mathrm{PGL}_2\ ,\ L_2(8) \longrightarrow G_2\ ,\ L_2(16) \longrightarrow D_8\ , L_2(32) \longrightarrow E_8\ ;$$
$$L_2(9) \longrightarrow \mathrm{PGL}_3\ ,\ L_2(27) \longrightarrow F_4.$$

Je ne sais pas en donner de description systématique.

Exemples

1) Si $G = \mathrm{SL}_2$, on a $r = 1$ et $d_1 = 2$; la condition dit alors que $q - 1$ divise 4, d'où $q = 3$ et $q = 5$, ce qui donne des plongements de $\mathrm{SL}_2(\mathbf{F}_3)$ et $\mathrm{SL}_2(\mathbf{F}_5)$ dans SL_2 ; d'où des plongements de $\mathrm{PSL}_2(\mathbf{F}_3) = \mathrm{Alt}_4$ et de $\mathrm{PSL}_2(\mathbf{F}_5) = \mathrm{Alt}_5$ dans PGL_2. On retrouve ainsi les groupes du tétraèdre et de l'icosaèdre (quant au groupe du cube, Sym_4, il s'interprète aussi comme $\mathrm{PGL}_2(\mathbf{F}_3)$, et c'est le normalisateur du groupe Alt_4).

2) Si $G = G_2$, on a $r = 2$, $d_1 = 2$, $d_2 = 6$; la condition dit que $q - 1$ divise 12, ce qui donne des plongements projectifs pour $q = 3, 5, 7, 13$. En fait, si $q = 7$ ou 13, ces plongements projectifs sont de vrais plongements de $L_2(q)$, car sinon leurs images seraient contenues dans le centralisateur d'un élément d'ordre 2, qui est de type $A_1 \cdot A_1$, et cela contredirait l'exemple 1. (Ce genre d'argument s'applique à beaucoup d'autres cas : les plongements projectifs intéressants sont de vrais plongements.)

3) Si $G = E_8$, on a $(d_1, \ldots, d_8) = (2, 8, 12, 14, 18, 20, 24, 30)$ et l'on en déduit notamment des plongements de $L_2(q)$ pour $q = 16, 31, 41, 49, 61$.

4) Le plus grand des entiers d_i est le *nombre de Coxeter* h, égal à $(\dim G)/r - 1$. L'énoncé 2.1.1 contient donc comme cas particulier la *conjecture de Kostant* : *si* $q = 2h + 1$ *est une puissance d'un nombre premier, le groupe* G^{ad} *contient un sous-groupe isomorphe à* $L_2(q)$.

5) On a un énoncé analogue à celui de Kostant lorsque $h + 1$ est une puissance d'un nombre premier, d'où par exemple $L_2(19) \to E_7^{\mathrm{ad}}$ et $L_2(31) \to E_8$. Lorsque $h+1$ est égal à un nombre premier p, on a un résultat plus précis (cf. [31]) : le groupe $\mathrm{PGL}_2(\mathbf{F}_p)$ (qui est "deux fois plus grand" que $\mathrm{PSL}_2(\mathbf{F}_p)$) est, lui aussi, plongeable dans G^{ad}. Lorsque $G = \mathrm{PGL}_2$, on retrouve le groupe du cube $\mathrm{PGL}_2(\mathbf{F}_3)$, cf. exemple 1. De ce point de vue, on peut dire que "les analogues" pour E_8 des groupes Alt_4, Sym_4, et Alt_5 du début de l'exposé sont respectivement $\mathrm{PSL}_2(\mathbf{F}_{31})$, $\mathrm{PGL}_2(\mathbf{F}_{31})$ et $\mathrm{PSL}_2(\mathbf{F}_{61})$.

2.2. Le critère 2.1.1 : démonstration de la nécessité

Il s'agit de prouver que, si $f : \mathrm{SL}_2(\mathbf{F}_q) \to G$ est un homomorphisme non dégénéré de type toral, alors $q - 1$ divise l'un des entiers $2d_i$.

On utilise :

2.2.1. *Soit* A *un* p*-sous-groupe élémentaire toral de* G*, et soit* $g \in N_G(A)$. *Soit* $I_g \in \mathrm{GL}(A)$ *l'automorphisme de* A (vu comme espace vectoriel sur $\mathbf{F}_p$) *défini par la conjugaison par* g. *Soit* λ *une valeur propre de* I_g *dans* $\overline{\mathbf{F}}_p$ *et soit* m *l'ordre de* λ (*dans* $\overline{\mathbf{F}}_p^*$). *Alors* m *divise l'un des* d_i.

(On peut supposer que A est contenu dans T ; d'après 1.1.1, il existe $w \in W$ qui induit I_g sur A. L'une des valeurs propres de w en caractéristique 0 a pour réduction λ en caractéristique p. Son ordre est donc de la forme mp^a, avec $a \geq 0$. D'après un théorème de Springer ([32], th. 3.4 (i)), mp^a divise l'un des d_i. Il en est donc de même de m.)

Revenons à f, et au p-Sylow $U = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$. Soit $A = f(U)$, et soit $g = f(h)$, où $h = \begin{pmatrix} c & 0 \\ 0 & c^{-1} \end{pmatrix}$ est un générateur du sous-groupe diagonal de $\mathrm{SL}_2(\mathbf{F}_q)$. Si l'on identifie U à $\mathbf{F}_q$, l'action de h sur ce groupe est l'homothétie de rapport c^2 ; ses valeurs propres (dans $\overline{\mathbf{F}}_p$) sont les conjugués de c^2, qui sont d'ordre $m = (q-1)/2$ si $p > 2$ et $m = q - 1$ si $p = 2$. En appliquant 2.2.1 à A et g, on voit que m divise l'un des entiers d_i ; donc $q - 1$ divise l'un des entiers $2d_i$.

Remarque.— Le même argument montre que, si f est un homomorphisme de $\mathrm{GL}_2(\mathbf{F}_q)$

dans G, qui est non dégénéré et de type toral (en un sens évident), alors $q-1$ divise l'un des d_i.

2.3. Le critère 2.1.1 : vérification de la suffisance

On doit montrer que, si $q-1$ divise l'un des entiers $2d_i$, il existe $f : \mathrm{SL}_2(\mathbf{F}_q) \to G$ qui est non dégénéré de type toral.

On ne connaît pas de démonstration générale de cet énoncé. On procède cas par cas :

1) Le cas où G est de type classique se traite facilement, grâce à la connaissance de la table des caractères de $\mathrm{SL}_2(\mathbf{F}_q)$; les caractères irréductibles de degré $(q \pm 1)/2$ sont particulièrement utiles. La condition de toralité est trivialement satisfaite si $p \neq 2$; dans le cas où $p = 2$, et où G est un groupe orthogonal, il faut faire un peu attention. La même méthode s'applique à G_2 (voir aussi [2], [11], [28]).

2) Pour les groupes exceptionnels, les inclusions des groupes classiques dans ceux-ci, et les plongements

$$G_2 \longrightarrow F_4 \longrightarrow E_6 \longrightarrow E_7 \longrightarrow E_8,$$

montrent qu'il suffit de traiter les cas suivants :

$$F_4\ (q = 25) \quad ; \quad E_6\ (q = 19) \quad ; \quad E_7\ (q = 29, 37) \quad ; \quad E_8\ (q = 16, 31, 41, 49, 61).$$

Le cas $(E_8\,;\,16)$ se traite en remarquant que $L_2(16)$ se plonge dans un groupe de type D_8, donc dans un groupe de type E_8 ; ce plongement n'est pas de type toral dans D_8, mais il le devient dans E_8 comme on le voit en appliquant [8], prop. 3.8.

Le cas $(F_4\,;\,25)$ se déduit de ce que $L_2(25)$ se plonge dans le groupe de Tits ${}^2F_4(2)'$, qui lui-même se plonge dans E_6, cf. 2.4.2, b) ci-après. On vérifie par un calcul de caractères que le sous-groupe de E_6 ainsi obtenu est contenu dans un conjugué de F_4, cf. Cohen–Wales [14].

Les cas $(E_6\,;\,19)$, $(E_7\,;\,37)$, $(E_8\,;\,31)$, $(E_8\,;\,41)$, $(E_8\,;\,49)$, $(E_8\,;\,61)$ ont été vérifiés par des calculs sur ordinateur, cf. Cohen–Wales [14], Kleidman–Ryba [24], Griess–Ryba [19] et [20], Cohen–Griess–Lisser [9].

Les cas $(E_7\,;\,37)$, $(E_8\,;\,31)$ et $(E_8\,;\,61)$ sont traités dans [31] par une méthode p-adique qui consiste à relever un plongement (bien choisi) de la caractéristique p à la caractéristique 0. Une variante non encore publiée de cette méthode permet de traiter aussi $(E_6\,;\,19)$, $(E_7\,;\,29)$ et $(E_8\,;\,41)$. Ainsi, *tous les cas où q est premier peuvent être obtenus sans ordinateur.*

Remarque.— Les calculs sur ordinateur ont un inconvénient évident : ils ne sont pas vérifiables pas à pas, comme une démonstration doit l'être. Ils ont toutefois un avantage : dans certains cas, ils montrent *l'unicité* du plongement considéré (à conjugaison près), cf. [19], [20]. C'est là un résultat que la méthode p-adique ne donne pas, au moins pour le moment.

2.4. Plongements projectifs des groupes finis simples dans les groupes de type exceptionnel

La table suivante est extraite de Griess-Ryba [22] (avec une petite correction relative à F_4). Elle donne la liste des groupes simples ayant un plongement projectif dans $G_2, \ldots, E_8$. Je renvoie à [22] et [27] pour divers renseignements supplémentaires sur ces plongements, ainsi que pour des références.

Table

$\mathbf{G_2}$ – Alt_n, $n = 5, 6$; $L_2(q)$, $q = 7, 8, 13$; $\mathrm{SU}_3(3) = G_2(2)'$.

$\mathbf{F_4}$ – *ceux de* G_2 *et* : Alt_n, $n = 7, 8, 9, 10$; $L_2(q)$, $q = 17, 25, 27$; $L_3(3)$; $\mathrm{SU}_4(2)$; $\mathrm{Sp}_6(2) = O_7(2)$; $O_8^+(2)$; ${}^3D_4(2)$.

$\mathbf{E_6}$ – *ceux de* F_4 *et* : Alt_{11} ; $L_2(q)$, $q = 11, 19$; $L_3(4)$; $\mathrm{PSU}_4(3)$; ${}^2F_4(2)'$; M_{11} ; $HJ = J_2$.

$\mathbf{E_7}$ – *ceux de* E_6 *et* : Alt_n, $n = 12, 13$; $L_2(q)$, $q = 29, 37$; $\mathrm{PSU}_3(8)$; M_{12}.

$\mathbf{E_8}$ – *ceux de* E_7 *et* : Alt_n, $n = 14, 15, 16, 17$; $L_2(q)$, $q = 16, 31, 32, 41, 49, 61$; $L_3(5)$; $\mathrm{PSp}_4(5)$; $G_2(3)$; ${}^2B_2(8) = Sz(8)$.

(Les notations sont celles de l'ATLAS [15]. En particulier $L_n(q)$ désigne le groupe $\mathrm{PSL}_n(\mathbf{F}_q)$. Vu que $\mathrm{Alt}_5 = L_2(4) = L_2(5)$ et $\mathrm{Alt}_6 = L_2(9)$, la liste pour G_2 pourrait aussi être écrite :

$\mathbf{G_2}$ – $L_2(q)$, $q = 4, 5, 7, 8, 9, 13$; $\mathrm{SU}_3(3) = G_2(2)'$.

De même, pour F_4, on peut remplacer Alt_8 par $L_4(2)$.)

La vérification de l'exactitude de cette table comporte deux parties. Tout d'abord :

2.4.1. *Un groupe simple qui ne figure pas dans la table n'a pas de plongement projectif dans* G.

Comme on peut s'y attendre, le point de départ est la *classification des groupes simples finis*, qui est admise (le lecteur curieux de savoir quelle partie de cette classification reste à démontrer pourra consulter Aschbacher [3]). Cela permet de passer

en revue les différents cas possibles : groupes alternés, groupes de type algébrique, groupes sporadiques. Pour éliminer un groupe S, on utilise des arguments variés, par exemple 1.2.2 ou 2.2.1 (qui suffisent si le groupe est très gros), ou (dans les cas difficiles) la table des caractères du groupe. C'est un travail délicat. La moindre erreur peut conduire à éliminer à tort le groupe en question. C'est ce qui s'était passé dans une liste précédente [8] pour les groupes $L_2(41)$, $L_2(49)$ et $Sz(8)$ qui avaient été déclarés non plongeables dans E_8.

2.4.2. *Tout groupe figurant dans la table a au moins un plongement projectif dans G.*

On utilise différentes méthodes. Par exemple :

a) Le cas le plus facile est celui où l'on connaît un sous-groupe de G dans lequel S a un plongement projectif. Ainsi, pour traiter le cas de Alt_{10} et F_4, il suffit de remarquer que Alt_{10} a une représentation orthogonale évidente de degré 9, autrement dit se plonge dans un groupe de type B_4, et l'on utilise le plongement de B_4 dans F_4.

b) Certains cas peuvent se traiter à partir de la table des caractères de S (et de ses extensions centrales). Outre $G = G_2$, déjà signalé, il faut mentionner le cas où $G = E_6$ et où S est le groupe de Tits ${}^2F_4(2)'$ (Cohen–Wales [14]). On part du fait que le groupe $S \cdot 2 = {}^2F_4(2)$ a une représentation irréductible V de dimension 78 (cf. [15], p. 75). Un calcul de caractères montre que $\wedge^2 V$ contient V ; il existe donc un homomorphisme non nul $\wedge^2 V \to V$ compatible avec l'action de $S \cdot 2$, et un autre calcul de caractères montre que l'identité de Jacobi est satisfaite. D'où une structure d'algèbre de Lie sur V. Il est clair que cette algèbre de Lie est simple ; puisqu'elle est de dimension 78, elle est de type B_6, C_6 ou E_6. On élimine les types B_6 et C_6 qui conduiraient à des représentations de $S \cdot 2$ de degré trop petit. L'algèbre de Lie V est donc de type E_6, ce qui fournit un plongement de $S \cdot 2$ dans E_6^{ad}, donc *a fortiori* un plongement de S. (On aimerait avoir davantage d'exemples de ce genre !)

c) La plupart des autres plongements ont été construits au moyen de calculs sur ordinateur. Je renvoie à [22] pour une description des méthodes employées. Je signale seulement que les calculs ne se font pas sur le corps k, mais sur un corps fini $\mathbf{F}_\ell$, où ℓ est un nombre premier ne divisant pas l'ordre de S et tel que $\mathbf{F}_\ell$ contienne les racines de l'unité intervenant dans la construction : ainsi, pour plonger $L_2(61)$ dans E_8, Cohen–Griess–Lisser [9] choisissent $\ell = 1831$. Le relèvement de $\mathbf{F}_\ell$ à $\mathbf{Z}_\ell$ (donc à la caractéristique 0) ne présente aucune difficulté vu que ℓ ne divise pas $|S|$. Il semble que, dans chaque cas, le calcul comporte suffisamment de vérifications internes pour qu'on puisse lui faire confiance.

2.5. Compléments

2.5.1. *Classification en caractéristique* > 0

L'analogue de 2.4 en caractéristique p a été fait par Liebeck–Seitz [27]. Tout groupe S intervenant en caractéristique 0 intervient aussi en caractéristique p (quel que soit p) ; c'est là une conséquence simple de la théorie de Bruhat–Tits, cf. [31], § 5. Outre ces groupes, et ceux qui sont "de caractéristique p", Liebeck–Seitz donnent la liste suivante :

$\mathbf{G_2}$ – $p = 2 : J_2$; $p = 5 : \mathrm{Alt}_7$; $p = 11 : J_1$.

$\mathbf{F_4}$ – *ceux de* G_2 *et* $p = 2 : L_4(3)$; $p = 3 : L_3(4)$; $p = 5 : Sz(8)$; $p = 11 : M_{11}$, Alt_{11}.

$\mathbf{E_6}$ – *ceux de* F_4 *et* $p = 2 : M_{12}$, Alt_{12}, $G_2(3)$, $O_7(3)$, M_{22}, J_3, Fi_{22} ; $p = 3 : M_{12}$, Alt_{12} ; $p = 5 : M_{12}$; $p = 7 : M_{22}$.

$\mathbf{E_7}$ – *ceux de* E_6 *et* $p = 5 : M_{22}$, Ru, HS ; $p = 7 : \mathrm{Alt}_{14}$.

$\mathbf{E_8}$ – *ceux de* E_7 *et* $p = 2 : L_4(5)$; $p = 3 : \mathrm{Alt}_{18}$, Th ; $p = 5 : Sz(32)$.

Noter en particulier le groupe de Janko J_1 dans $G_2(\mathbf{F}_{11})$ et le groupe de Thompson Th dans $E_8(\mathbf{F}_3)$.

2.5.2. *Classes de conjugaison de plongements*

On aimerait pouvoir compléter la table 2.4 en décrivant les plongements à conjugaison près. Cela a été fait dans certains cas, mais pas dans tous, cf. [22]. Le cas des plongements de Alt_5 dans E_8 est particulièrement intéressant (cf. Frey [16]) ; on peut déterminer les triplets (x, y, z) de classes de conjugaison de E_8 d'ordres $(2, 3, 5)$ qui sont représentables dans un même sous-groupe Alt_5. Pour tous ces triplets, sauf un (celui appelé "844" dans [16]), Frey détermine le nombre de classes de conjugaison correspondantes (une ou deux). Par contre, pour le cas "844" (qui est le seul où le centralisateur du sous-groupe Alt_5 soit fini), on ne sait pas combien il y a de classes de conjugaison ; on dispose de plusieurs tels sous-groupes (par exemple un sous-groupe du groupe de Borovik [7], ou un sous-groupe de $L_2(41)$, ou de $L_2(61)$,...), mais il n'est pas facile de voir s'ils sont ou non conjugués. Comme Alt_5 admet la présentation :

$$(x, y, z \mid x^2 = y^3 = z^5 = 1 \ , \ xyz = 1),$$

c'est là un problème analogue à celui de la "rigidité" intervenant pour la classification des revêtements galoisiens de la droite projective ramifiés en 3 points.

2.5.3. *Rationalité*

On sait que G provient par extension des scalaires d'un groupe *déployé* G_{dep} défini sur $\mathbf{Q}$. Si S (ou $\overline{S}$) est plongeable dans $G(k)$, on peut se demander quels sont les sous-corps k' de k tels que S soit plongeable dans $G_{\mathrm{dep}}(k')$. Cette question est étroitement liée à la précédente (celle des classes de conjugaison) : voir là-dessus [19], App. 2. Voici un exemple typique :

D'après Aschbacher [2], le groupe $S \cdot 2 = G_2(2)$ admet un plongement dans $G_2(k)$, et un seul, à conjugaison près. Or, à la fois $S \cdot 2$ et G_2 ont un centre trivial, et pas d'automorphisme externe. De plus, le centralisateur de $2 \cdot S$ est trivial. Soit P l'ensemble de ces plongements ; c'est un G_2-torseur qui est défini de façon naturelle sur $\mathbf{Q}$. Il définit donc une $\mathbf{Q}$-forme G_2^0 de G_2, et l'on peut plonger $S \cdot 2$ dans $G_2^0(\mathbf{Q})$ par définition même de G_2^0. Or, il n'y a que deux formes de G_2 sur $\mathbf{Q}$, que l'on distingue par leurs points réels ; la forme déployée ne peut pas contenir $2 \cdot S$: son compact maximal est trop petit. Ainsi, G_2^0 est la forme non déployée de G_2, celle qui correspond aux octonions usuels. On conclut de là que le plongement cherché de $S \cdot 2$ dans $G_{\mathrm{dep}}(k')$ existe si et seulement si G_{dep} et G_2^0 sont k'-isomorphes, i.e. *si et seulement si* -1 *est somme de* 4 *carrés dans* k'. Un argument analogue montre que $L_2(13)$ est plongeable dans $G_{\mathrm{dep}}(k')$ si et seulement si k' contient $\sqrt{13}$ et -1 est somme de 4 carrés dans k' ; même chose pour $L_2(8)$, avec $\sqrt{13}$ remplacé par $z_9+\overline{z}_9$, où z_9 est une racine primitive 9-ème de l'unité. (Noter l'analogie de ces énoncés avec le suivant, connu depuis longtemps : Alt_4, Sym_4 et Alt_5 sont plongeables dans $\mathrm{PGL}_2(k')$ si et seulement si -1 est somme de 2 carrés dans k' et (pour Alt_5) k' contient $\sqrt{5}$.)

BIBLIOGRAPHIE

[1] J.F. ADAMS - 2-*tori in* E_8, Math. Ann. **287** (1987), 29-39 (= *Selected Works*, vol. II, 264-274).

[2] M. ASCHBACHER - *Chevalley groups of type* G_2 *as the group of a trilinear form*, J. Alg. **109** (1987), 193-259.

[3] M. ASCHBACHER - *Quasithin groups*, in *Algebraic Groups and their Representations* (R. Carter and J. Saxl edit.), NATO AS series, vol. **517**, 321-340, Kluwer, 1998.

[4] A. BOREL - *Sous-groupes commutatifs et torsion des groupes de Lie compacts connexes*, Tôhoku Math. J. **13** (1961), 216-240 (= *Oe.* II, n° 53 et *Commentaires*, 775-777).

[5] A. BOREL and G.D. MOSTOW - *On semi-simple automorphisms of Lie algebras*, Ann. Math. **61** (1955), 389-405 (= A. Borel, *Oe.* I, n° 36).

[6] A. BOREL et J.-P. SERRE - *Sur certains sous-groupes des groupes de Lie compacts*, Comm. Math. Helv. **27** (1953), 128-139 (= A. Borel, *Oe.* I, n° 24).

[7] A.V. BOROVIK - *A maximal subgroup in the simple finite group* $E_8(q)$, Contemp. Math. A.M.S. **131** (1992), vol. I, 67-79.

[8] A.M. COHEN and R.L. GRIESS, Jr - *On finite simple subgroups of the complex Lie groups of type* E_8, AMS Proc. Symp. Pure Math. **47** (1987), vol. II, 367-405.

[9] A.M. COHEN, R.L. GRIESS, Jr and B. LISSER - *The group* $L(2,61)$ *embeds in the Lie group of type* E_8, Comm. Alg. **21** (1993), 1889-1907.

[10] A.M. COHEN and G.M. SEITZ - *The r-rank of the groups of exceptional Lie type*, Indag. Math. **49** (1987), 251-259.

[11] A.M. COHEN and D.B. WALES - *Finite subgroups of* $G_2(\mathbf{C})$, Comm. Alg. **11** (1983), 441-459.

[12] A.M. COHEN and D.B. WALES - *Embeddings of the group* $L(2,13)$ *in groups of Lie type* E_6, Israel J. Math. **82** (1993), 45-86.

[13] A.M. COHEN and D.B. WALES - *Finite simple subgroups of semisimple complex Lie groups – a survey*, in *Groups of Lie type and their geometries*, LMS Lect. Notes **207** (1995), 77-96.

[14] A.M. COHEN and D.B. WALES - *Finite subgroups of* $F_4(\mathbf{C})$ *and* $E_6(\mathbf{C})$, Proc. London Math. Soc. **74** (1997), 105-150.

[15] J.H. CONWAY, R.T. CURTIS, S.P. NORTON, R.A. PARKER and R.A. WILSON - *Atlas of Finite Groups*, Clarendon Press, Oxford, 1985.

[16] D. FREY - *Conjugacy of alternating groups of degree* 5 *and* $\mathrm{SL}(2,5)$ *subgroups of the complex Lie groups of type* E_8, Memoirs AMS **634** (1998).

[17] R.L. GRIESS, Jr - *Elementary abelian subgroups of algebraic groups*, Geom. Dedicata **39** (1991), 253-305.

[18] R.L. GRIESS, Jr and A.J.E. RYBA - *Embeddings of* $U(3,8)$, $Sz(8)$ *and the Rudvalis group in algebraic groups of type* E_7, Invent. math. **116** (1994), 215-241.

[19] R.L. GRIESS, Jr and A.J.E. RYBA - *Embeddings of* $\mathrm{PGL}(2,31)$ *and* $\mathrm{SL}(2,32)$ *in* $E_8(\mathbf{C})$, Duke Math. J. **94** (1998), 181-211.

[20] R.L. GRIESS, Jr and A.J.E. RYBA - *Embeddings of* $\mathrm{PSL}(2,41)$ *and* $\mathrm{PSL}(2,49)$ *in* $E_8(\mathbf{C})$, J. Symb. Comp. **11** (1999), 1-17.

[21] R.L. GRIESS, Jr and A.J.E. RYBA - *Embeddings of* $Sz(8)$ *into exceptional Lie groups*, soumis au J. Crelle.

[22] R.L. GRIESS, Jr and A.J.E. RYBA - *Finite simple groups which projectively embed in an exceptional Lie group are classified !*, Bull. AMS **36** (1999), 75-93.

[23] A. GROTHENDIECK - *Torsion homologique et sections rationnelles*, Sém. Chevalley (1958), *Anneaux de Chow et Applications*, exposé 5.

[24] P.B. KLEIDMAN and A.J.E. RYBA - *Kostant's conjecture holds for* E_7 : $L_2(37) < E_7(\mathbf{C})$, J. Alg. **161** (1993), 535-540.

[25] M.-A. KNUS, A. MERKURJEV, M. ROST and J.-P. TIGNOL - *The Book of Involutions*, AMS Colloquium Publ. **44**, 1998.

[26] M.W. LIEBECK - *Subgroups of exceptional groups*, in *Algebraic Groups and their Representations* (R.W. Carter and J. Saxl edit.), NATO AS series, vol. **517**, 275-290, Kluwer, 1998.

[27] M.W. LIEBECK and G.M. SEITZ - *On finite subgroups of exceptional algebraic groups*, 1999, à paraître au J. Crelle.

[28] A. MEURMAN - *An embedding of* PSL$(2,13)$ *in* $G_2(\mathbf{C})$, Lect. Notes in Math. **933** (1982), 157-162.

[29] Z. REICHSTEIN and B. YOUSSIN - *Essential dimensions of algebraic groups and a resolution theorem for G-varieties*, 1999, à paraître.

[30] J.-P. SERRE - *Cohomologie galoisienne : progrès et problèmes*, Sém. Bourbaki 1993-94, exposé n° 783 (SMF, Astérisque **227** (1995), 229-257).

[31] J.-P. SERRE - *Exemples de plongements des groupes* $\mathbf{PSL}_2(\mathbf{F}_p)$ *dans des groupes de Lie simples*, Invent. math. **124** (1996), 525-562.

[32] T.A. SPRINGER - *Regular elements of finite reflection groups*, Invent. math. **25** (1974), 159-198.

[33] T.A. SPRINGER and R. STEINBERG - *Conjugacy classes*, Lect. Notes in Math. **131** (1970), 281-312 (= R. Steinberg, *C.P.*, 293-323).

[34] R. STEINBERG - *Torsion in reductive groups*, Adv. in Math. **15** (1975), 63-92 (= *C.P.*, 415-444).

[35] J. TITS - *Normalisateurs de tores*, I. *Groupes de Coxeter étendus*, J. Alg. **4** (1966), 96-116.

[36] J. TITS - *Sur les constantes de structure et le théorème d'existence des algèbres de Lie semi-simples*, Publ. Math. IHES **31** (1966), 21-58.

Jean-Pierre SERRE

Collège de France
3, rue d'Ulm
F-75005 PARIS
E-mail : serre@dmi.ens.fr